DOCUMENT RESUME

ED 291 248 FL 017 176

AUTHOR	Harley, Birgit; And Others
TITLE	The Development of Bilingual Proficiency. Final Report. Volume I: The Nature of Language Proficiency, Volume II: Classroom Treatment, Volume III: Social Context and Age.
INSTITUTION	Ontario Inst. for Studies in Education, Toronto. Modern Language Centre.
PUB DATE	Apr 87
NOTE	841p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF05/PC34 Plus Postage.
DESCRIPTORS	*Bilingualism; Classroom Observation Techniques; *Classroom Techniques; Elementary Secondary Education; English (Second Language); Foreign Countries; French; *Immersion Programs; Instructional Effectiveness; Japanese; *Language Acquisition; Language Patterns; *Language Proficiency; Language Usage; Longitudinal Studies; Native Speakers; Portuguese; Program Effectiveness; Second Language Learning; Uncommonly Taught Languages; Urban Areas
IDENTIFIERS	*Ontario (Toronto)

ABSTRACT

The Development of Bilingual Proficiency is a large-scale, five-year research project begun in 1981. The final report contains three volumes, each concentrating on specific issues investigated in the research: (1) the nature of language proficiency, including second language lexical proficiency and the development and growth of metaphor comprehension in first and second languages; (2) the effect of classroom treatment of language proficiency, including a study of the Colt Observation Scheme, a core French observation study, an immersion observation study, and a classroom experiment in functional grammar in French immersion; and (3) the relevance of social context and age for language learning, including studies on the language use, attitudes and proficiency of Toronto Portuguese-Canadian children, development of bilingual proficiency and the transition from home to school in Portuguese-speaking children, the relationship of arrival age, residence length, and literacy skills among Japanese immigrant students, and the relationship between starting age and oral second language proficiency in three groups of classroom learners. Each volume contains a 20-page overview of the project and brief summaries of each individual study. (MSE)

****
*   Reproductions supplied by EDRS are the best that can be made   *
*   from the original document.   *
****

ED291248

THE DEVELOPMENT OF BILINGUAL PROFICIENCY

FINAL REPORT

Volume I: The Nature of Language Proficiency

Birgit Harley
Patrick Allen
Jim Cummins
Merrill Swain

April 1987

Modern Language Centre
The Ontario Institute for Studies in Education

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M Swain

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

FL017176

BEST COPY AVAILABLE

TABLE OF CONTENTS

Page

VOLUME I

Preface vii
Project Staff viii

Introduction
THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. Purpose 1
2. The Nature of Language Proficiency 1
3. Classroom Treatment Studies 8
4. Social Context and Age 13
5. Conclusions 19

Chapter 1
PERSPECTIVES ON LEXICAL PROFICIENCY IN A SECOND LANGUAGE

1. Introduction 23
2. The Data Base 24
3. Measures of Lexical Proficiency 26
4. Lexical Measures and L2 Proficiency 30
5. Lexical Use of Immersion Students and Native Speakers 34
6. Conclusion 40

Tables 44
Figures 60
Appendix A 64
Appendix B 66

Chapter 2
THE DEVELOPMENTAL GROWTH OF METAPHOR COMPREHENSION IN CHILDREN'S FIRST AND SECOND LANGUAGE

1. The Development of Bilingualism 67
2. A Cognitive Developmental Perspective 69
3. Summary of the Present Study and Hypotheses 70
4. Method 71
5. Results: Metaphors Interpreted in English 76
6. Results: Metaphor Interpretations in Spanish 81
7. Conclusions 84

Tables 91
Figures 101
Appendix 107

Page

VOLUME II
Preface vii
Project Staff viii

Introduction
THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. Purpose 1
2. The Nature of Language Proficiency 1
3. Classroom Treatment Studies 8
4. Social Context and Age 13
5. Conclusions 19

Chapter 3
COLT OBSERVATION SCHEME: DEVELOPMENT AND VALIDATION

1. Introduction 23
2. Description and Rationale of the Observation Scheme 24
3. The Validation Study 36
4. Discussion 43

Tables 50
Figures 52
Appendix A 54

Chapter 4
THE CORE FRENCH OBSERVATION STUDY

1. Introduction 56
2. The Sample 57
3. General Procedures 59
4. The Colt Observation Scheme: Procedures 61
5. Classroom Practice: Findings 64
6. Proficiency Predictions 70
7. Proficiency Tests 72
8. Process/Product Findings 78
9. Discussion and Interpretation of Results 86

Tables 98
Appendix A 127
Appendix B 137
Appendix C 161
Appendix D 164
Appendix E 167
Appendix F 181

Page

VOLUME II (continued)

Chapter 5
THE IMMERSION OBSERVATION STUDY

1. Introduction 190
2. Vocabulary Instruction in Immersion Classes 192
3. Vous/Tu Input 222
4. Student Talk in Teacher-Fronted Activities 226
5. Error Treatment 233
6. Concluding Comment 240

Tables 252
Appendix A 264
Appendix B 287
Appendix C 293
Appendix D 317

Chapter 6
FUNCTIONAL GRAMMAR IN FRENCH IMMERSION: A CLASSROOM EXPERIMENT

1. Purpose and Rationale 342
2. Design of the Experiment 346
3. The Findings 354
4. Discussion 361
5. Conclusion 363

Tables 367
Figures 379
Appendix A 381
Appendix B 385
Appendix C 386
Appendix D 389
Appendix E 390
Appendix F 391
Appendix G 399
Appendix H 409

Page

VOLUME III

Preface vii

Project Staff viii

Introduction

THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. Purpose 1
2. The Nature of Language Proficiency 1
3. Classroom Treatment Studies 8
4. Social Context and Age 13
5. Conclusions 19

Chapter 7

THE LANGUAGE USE PATTERNS, LANGUAGE ATTITUDES, AND BILINGUAL PROFICIENCY OF PORTUGUESE CANADIAN CHILDREN IN TORONTO

1. Introduction 23
2. Scoring Procedures and Refinement of the Data Set 28
3. Results 35
4. Comparison of Portuguese Proficiency with Azorean Native Speakers 49

Tables 54

Appendix A 85

Appendix B 88

Appendix C 106

Appendix D 119

Appendix E 154

Chapter 8

THE DEVELOPMENT OF BILINGUAL PROFICIENCY IN THE TRANSITION FROM HOME TO SCHOOL: A LONGITUDINAL STUDY OF PORTUGUESE-SPEAKING CHILDREN

1. Purpose of the Study 159
2. Methodology 159
3. Scoring Procedures and Preliminary Results for Home Interaction Data 165
4. Home Interaction of 4-Year Old Portuguese Background Children in Toronto: Preliminary Trends 176

Tables 182

Page

VOLUME III (continued)

Chapter 9
AGE OF ARRIVAL, LENGTH OF RESIDENCE, AND INTERDEPENDENCE OF LITERACY SKILLS AMONG JAPANESE IMMIGRANT STUDENTS

1. Background 183
2. Methodology 185
3. Results 188
4. Conclusion 191

Tables 194
Figures 199
Appendix A 200

Chapter 10
THE RELATIONSHIP BETWEEN STARTING AGE AND ORAL SECOND LANGUAGE PROFICIENCY IN THREE GROUPS OF CLASSROOM LEARNERS

1. Introduction 203
2. The Sample 204
3. Procedures 205
4. Verb Use: Analysis and Results 207
5. Fluency: Analysis and Results 213
6. Sociolinguistic Test Results 216
7. Discussion 216
8. Conclusion 220

Tables 223
Appendix A 232
Appendix B 235
Appendix C 253
Appendix D 251

PREFACE

The Development of Bilingual Proficiency is a large-scale, five-year research project which began in September 1981. The present Final Report of the project is the third in a series. It follows an interim Year 1 Report, produced in September 1982 at the end of the first year of the project, and a Year 2 Report, produced in September 1983.

There are three volumes in this Final Report of the project, each concentrating on specific issues investigated in the research: the nature of language proficiency (Volume I), the effect of classroom treatment on language proficiency (Volume II), and the relevance of social context and age for language learning (Volume III). Each volume is introduced by an identical 20-page overview of all the studies carried out in the context of the Development of Bilingual Proficiency (DBP) Project. The overview includes brief summaries of the individual studies together with an indication as to where the complete report of each study is to be found (either in the Year 2 Report or in Volume I, II, or III of the Final Report). Within the complete reports of individual studies contained in this Final Report, references to other Development of Bilingual Proficiency Project studies appear either as 'Year 2 Report' or, when they form part of the Final Report, as chapter numbers only. Note that Chapters 1 and 2 appear in Volume I, Chapters 3-6 in Volume II, and Chapters 7-10 in Volume III.

The authors wish to acknowledge the contribution of the many individuals and organizations who have played a role in the Development of Bilingual Proficiency Project since it began in September 1981. We are grateful to administrators, teachers, students and their parents in the following Ontario school boards for their participation in the research: the Board of Education for the City of Scarborough, the Carleton Board of Education, the Metropolitan Separate School Board, the North York Board of Education, and the Toronto Board of Education. In addition, we would like to thank the Portuguese Secretary of State for Immigration, the Regional Secretary of Social Affairs for the Autonomous Region of the Azores, and the staff, parents and students of the Japanese School of Toronto Shokokai Inc. We wish to acknowledge the contribution of Ellen Bialystok and Raymond Mougeon, who were principal investigators of the project in 1981-82 and 1981-83 respectively. We would also like to express our appreciation to project staff for their part in carrying out the research and in text-processing. Finally, we wish to acknowledge the financial support provided in the form of a five-year negotiated grant (No. 431-79-0003) by the Social Sciences and Humanities Research Council of Canada, and the administrative and financial contribution of the Ontario Institute for Studies in Education.

B.H., P.A., J.C., M.S.

PROJECT STAFF

(from September 1983)*

Principal Investigators:

Patrick Allen, Jim Cummins, Birgit Harley, Merrill Swain

Project Managers:

Birgit Harley, Sharon Lapkin

Research Officers:

Margarida Aguiar, Jud Burtis, Susanne Carroll, Maria Fröhlich, Vince Gaudino, Mary Lou King, Laurette Lévy, José Lopes, Françoise Pelletier, Bruno Piché, Jorge Ramos, Mireille Tremblay, Virginia Won (Pang)

Graduate Assistants:

Jane Appelt, Eva Bild, Patrick Conteh, Giselle Corbeil, Vera Held, Richard Kruk, Bill Marshall, Josée Mazzilli, Anne Riddle, Emilia Rivas-Arraiz, Kathryn Shred, Nina Spada, Yilin Sun, Tim Tomlinson

Post-doctoral Fellows:

Monica Heller, Janice Johnson

Statistical Consultant:

Jamshid Etezadi-Amoli

Text-processing Specialist:

Ellen Jeske

Secretaries:

Joyce Howlett, Doreen Panowyk, Donna Schillaci

---

* For staff members in previous years, see earlier reports

Introduction

THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. PURPOSE

The purpose of this five-year research project has been to investigate issues concerning language proficiency[1] and its development in educational contexts for children learning a second language. The research has concentrated on the following major issues: the nature of language proficiency; the impact of instructional practices on language learning; the relationship between social-environmental factors and bilingual proficiency; and the relationship between age and language proficiency. In this overview of the project, studies focussing on each of these issues are summarized.

2. THE NATURE OF LANGUAGE PROFICIENCY

The focus and ultimate goal of all studies carried out within the Development of Bilingual Proficiency Project is the improvement of educational practices as they relate to second language learning and teaching. Because so much of school practice relates rather narrowly to the teaching and learning of grammatical aspects of the target language, it was considered essential to broaden the scope of the typical educational definition of language proficiency to incorporate discourse and sociolinguistic dimensions, and to consider the differential demands that context-reduced versus more context-embedded language tasks may make on the learner.

2:1 Large-scale Proficiency Study
(Year 2 Report)

Objectives. The primary purpose of the large-scale proficiency study conducted during the first two years of the Project was to determine whether the three hypothesized traits, representing key components of language proficiency, could be empirically distinguished. It was hypothesized that grammatical, discourse, and sociolinguistic competence would emerge as distinct components of second language proficiency which may be differentially manifested under different task conditions. A secondary purpose of the study was to develop a set of exemplary test items and scoring procedures that could be used, or modified for use, in further studies involving the measurement of the hypothesized traits. A final purpose of the study was to provide a broadly based description of the target language proficiency of the second language learners tested, in relation to that of native speakers.

Subjects. A total of 198 students was involved in the study. Of these, 175 were grade 6 early French immersion students from the Ottawa region, and 23 were grade 6 native speakers from a regular Francophone school in Montreal. The immersion students, in six intact classes, had received 100% of their schooling in French in kindergarten to grade 2 or 3, since when they had been taught in English for a gradually increasing portion of each day. At the time of testing, about 50% of their school subjects were

being taught in French, and the other 50% in English. This sample of classroom second language learners was selected because of the theoretically interesting and educationally innovative nature of their intensive school-based language learning experience, and because they were at an age where they were sufficiently proficient in the second language to be able to cope with a wide range of types of language tasks.

Instruments. A multi-method multi-trait design was used to determine the extent to which grammatical, discourse, and sociolinguistic dimensions of the immersion students' French proficiency were distinguishable. To measure proficiency on each trait, three methods of testing were used: oral production, multiple choice, and written production. A matrix with nine test cells was thus created, consisting of three tests of grammar, three of discourse, and three of sociolinguistics. The oral production task for each trait was administered to a randomly selected sub-sample of 69 immersion students and ten native speakers, representing ten-eleven subjects from each class.

Grammatical competence was operationalized for the purposes of this study as rules of morphology and syntax, with a major emphasis on verbs and prepositions. The grammar oral production task consisted of a guided individual interview in which the interviewers' questions were designed to elicit a variety of verb forms and prepositions in French, as well as responses that were sufficiently elaborated to score for syntactic accuracy. The content of the interview questions (e.g. favourite pastimes, trips taken) was at the same time designed to focus the subject's attention on communication rather than the code. Grammatical scoring was based on the student's ability to use certain grammatical forms accurately in the context of particular questions. The group-administered grammar multiple choice test consisted of 45 written items which also assessed knowledge of the verb system, prepositions, and other syntactic rules, including written agreement rules. The student's task was to select the correct response from three alternatives provided. The third grammar task, written production, consisted of two short compositions to be written in 15 minutes each — one a narrative and the other a letter of request. Both this written production task and a parallel discourse written production task — also involving a narrative and a request letter — were assessed for grammatical proficiency. Scoring focussed on grammatical accuracy in verbs, prepositions, and other rules of syntax and morphology.

The discourse trait was defined as the ability to produce and recognize coherent and cohesive text (written or oral). For the individual discourse oral production task, the student was required to retell the story of a silent movie and to present arguments in support of an opinion. This task was rated on 5-point scales both globally and in detail for coherence and cohesion, focussing, for example, on the student's ability to make clear and accurate reference to characters, objects, and locations, to produce a logically connected text, and to fulfill the basic task required. The discourse multiple choice task consisted of 29 short written passages from each of which a sentence had been omitted. The student was required to select from three alternatives the sentence that best fit the context. The discourse written production task, like the grammar written production task, consisted of a narrative and a request letter. All four (grammar and discourse) tasks were rated for proficiency in discourse on the same kinds of features that were assessed in the discourse oral production task.

Sociolinguistic competence, the third trait dealt with in this proficiency study, was operationalized as the ability to produce and recognize socially appropriate language in context. The individual oral production task involved a set of slides with taped descriptions representing situations of different levels of formality. The student's task

was to respond appropriately with a request, offer, or complaint. Scoring focussed on the student's ability to shift register appropriately. Thus sociolinguistic proficiency was measured by difference scores, calculated by subtracting the number of formal 'politeness' markers produced by the student in informal variants of situations from those produced in formal variants of the situations. The sociolinguistic multiple choice test consisted of 28 items, each with three alternative ways of expressing a given sociocultural function. The choices were all grammatically accurate but not equally appropriate. The student's task was to select the most appropriate of the choices in the given situation. Scoring of responses was weighted according to the choices made by native speakers. The sociolinguistic written production task involved the writing of a formal request letter and two informal notes, all of which could be categorized as directives. The request letter written as part of the discourse written production task was also scored for sociolinguistic proficiency. As for the oral production task, scoring was based on difference scores, calculated by subtracting the number of formal markers produced in the notes from those produced in the letters.

Reliability and generalizability of scores. The component within-test scores were combined to produce a single overall score for each of the nine trait-method cells in the matrix. The composition of each of these overall scores was calculated to maximize validity and reliability. On the multiple choice tests, the reliability of the immersion students' total scores ranged from .58 on the sociolinguistic test to .75 on the discourse test. Generalizability studies were conducted on those cells for which sufficient data were available: the sociolinguistic oral production test and the three written production tests. G-coefficients for these tests, based on the subsample of orally tested students, were comparable to the multiple choice test reliabilities.

Testing a model of proficiency. In order to determine whether the three traits -- grammatical, discourse, and sociolinguistic competence -- could be empirically distinguished, two kinds of analyses were performed: (a) factor analysis, and (b) a comparison of the group means of the learners and native speakers.

The factor analysis based on the 69 orally tested immersion students failed to confirm the hypothesized three-trait structure of proficiency. Instead, confirmatory factor analysis by means of LISREL produced a two-factor solution. One of these factors was interpretable as a general language proficiency factor; it had positive loadings from all cells in the nine-test matrix except for the sociolinguistic written production test. The highest loadings on this general factor were from the three grammatical tests. The second factor was interpretable as a written method factor; it had loadings from the three multiple choice tests and from all three written production tests. The tests loading on this method factor appeared to be tapping the kind of literacy-oriented linguistic proficiency that is typically learned in classrooms. The lack of trait structure emerging from the factor analysis may have been due to the homogeneity of the immersion student sample. Within their classroom setting these students had all had very much the same kind of exposure to French, and strong opportunities for some students to develop proficiency in one area, and other students to develop proficiency in a different area, were lacking.

A different kind of result emerged from comparisons of immersion and native-speaker scores on the various tests. On all three grammar tests, the immersion students' mean score was considerably lower than that of the native speakers (p .01), and they also scored generally lower on the sociolinguistic tests than did the native speakers. On the discourse tasks, however, the scores of the immersion students were close or

equivalent to those of the native speakers and there were no significant between-group differences. The immersion students' strong performance in discourse may have been due to positive transfer from prior experience in their mother tongue. In contrast to the factor analysis results, then, these comparative findings, showing very different results for discourse as opposed to grammar and sociolinguistics, provide some evidence in support of a distinction between traits.

Conclusions. It was concluded that, although the three hypothesized language proficiency traits were not empirically distinguished via the factor analysis, this result may have been dependent on the relatively homogeneous language learning background of the immersion population studied. This did not necessarily mean that the traits would not be distinguishable in a more heterogeneous language learning population. From an educational perspective it was clear that the analysis of proficiency into different components was diagnostically revealing of the second language strengths and weaknesses of the immersion students. It was decided that two kinds of further studies were indicated to probe issues concerning how different dimensions of proficiency develop as a function of the immersion students' specific language learning experience: (a) small-scale in-depth studies of specific aspects of the immersion students' second language proficiency based on the data already collected (see 2:2 - 2:3 below), and (b) the study of language learning activities in the immersion classroom setting (see 3:3 - 3:4 below).

2:2 Transfer in Immersion Students' Compositions
(Year 2 Report)

Hypotheses and design. Given the shared mother tongue, English, of the immersion students and the dominance of English in the wider school and outside-school environment of the immersion program, mother tongue transfer was expected to be a continuing factor in the students' written production at the grade 6 level. In a small-scale study of compositions written by 22 native speakers and 22 of the orally tested immersion students from two randomly selected classes in the larger proficiency study, evidence was sought for the hypothesis that mother tongue transfer may be manifested in the way in which the learners were distributing semantic information across syntactic elements in the second language, without necessarily making outright errors.

One of the composition topics assigned in the large-scale proficiency study, Au secours!, involved writing a story about the rescue of a kitten from a tree. The students' stories on this topic contained a very similar series of events, involving several changes of location. The focus of the present study was on how the immersion students were expressing the location/direction distinction in these stories, given that there are characteristic differences between French and English in this linguistic domain. While in English, prepositions generally serve an important role in conveying the location/direction distinction (e.g. at/to, in/into), in French there is a general tendency for direction to be expressed in the verb, and for prepositions (e.g. à, dans, sur) to be neutral with respect to the location/direction distinction. It was hypothesized that the immersion students would rely on prepositions rather than the verb to express the notion of direction.

Findings. A comparison of directional expressions in the Au secours! stories written by the immersion students and the native speakers showed that, as expected, the immersion students were much less likely than the native speakers to mark direction in the verb, preferring a non-directional verb of motion such as courir together with a

preposition to express the directional notion. The immersion students, at the same time, sometimes erroneously used French prepositions unmarked for direction as if they were carrying the directional distinction, and also tended to make more use than the native speakers of prepositional phrases expressing direction, even on those occasions when they also used directional verbs. This latter tendency did not necessarily lead to error. The findings of the study thus provide support for the hypothesis that the immersion students would show a systematic tendency to rely more heavily on prepositions to express the notion of direction than the native speakers.

Conclusions. It was concluded that the students may need more focussed classroom input that would alert them to such characteristics of French that are different from English, together with more opportunities for expressing the relevant distinctions in their second language.

2:3 Lexical Proficiency in a Second Language (Final Report, Vol. I)

In the large-scale proficiency study described above (2:1), there were no measures specifically designed to assess lexical proficiency, not because lexical proficiency was considered unimportant but because it was assumed to enter into performance on all the tasks assigned. In the present study, the two narratives and three request letters written by 69 immersion students and 22 native speakers in the context of the various written production tests were re-analysed from a lexical perspective, with verbs being selected as the focus for the study. The purpose of the study was threefold: (a) to compare different quantitative measures of immersion students' lexical proficiency in their second language (L2) writing; (b) to examine the relationship between written lexical proficiency and other aspects of their L2 communicative competence, and (c) to describe the students' lexical use in relation to that of native speakers.

(a) Measures of lexical proficiency. Five quantitative measures of lexical proficiency were developed and statistically compared. One of these was a 'lexical error rate', while the other four were variations on the theme of lexical richness, labelled respectively 'number of lexical types', 'lexical variety', 'lexical specificity', and 'lexical sophistication'. All the measures, except for 'number of lexical types' were controlled for length of text. For each student the data from the five written compositions were lumped together. Two of the relatively difficult measures were retained as the most appropriate for further use in a factor analysis. The first was 'lexical specificity', which consisted of the number of different verb types used by each student, not counting the 20 most frequent verbs in French or those that were used in the instructions to the compositions, divided by the number of verb items produced. The second measure was 'lexical sophistication', representing those relatively infrequent verbs not found in a basic word frequency list, also divided by the number of verb items produced.

(b) Lexical measures and L2 proficiency. Three mutually exclusive hypotheses arising from previous work were examined via factor analysis: (1) that lexical proficiency is equally involved in all three of the components of language proficiency examined in the large-scale proficiency study: namely, grammar, discourse, and sociolinguistics; (2) that lexical proficiency is part of the grammar component; or (3) that lexical proficiency is a separate component, distinct from the other three components of language proficiency.

Confirmatory factor analyses showed that an acceptable solution to fit any one of these three hypotheses could be found, and that there was no conclusive evidence favouring any one of the three hypotheses over the other two. One interesting finding was that in the three- and four-factor solutions corresponding to hypotheses (2) and (3) respectively, a grammar factor and a discourse factor emerged, which had not previously been found in the large-scale study where no lexical measures had been included.

(c) Lexical use of immersion students and native speakers. A comparison of the verb lexis used by the immersion students and the native speakers in their compositions revealed that the immersion students tended to make proportionately more use of high-coverage verbs, and less use of some morphologically or syntactically complex verbs such as pronominal and derived verbs. The inflectional complexity of some high coverage verbs did not appear to be a deterrent to their use although inflectional errors (considered grammatical rather than lexical errors) did occur. Semantic and syntactic incongruence with their English mother tongue (L1) emerged as an important factor in the immersion students' non-use of some French verb types and in the lexical errors they made. At the same time, the students demonstrated positive L1 transfer in the use of some cognate verbs in French.

Conclusions. It was suggested that the immersion students' stock of lexical items would benefit from more classroom activities designed to increase their use of L2 derivational resources and to emphasize the use of more specific vocabulary.

2:4 Communicative Skills of Young L2 Learners (Year 2 Report)

Purpose and data base. This exploratory study involved a detailed investigation of methods of scoring oral L2 performance and of the interrelationships among various aspects of L2 proficiency. The study was based on a subset of data previously collected in the context of another Modern Language Centre project. It consisted of oral tasks in English with 22 Japanese immigrant students in grades 2, 3, 5 and 6, together with academic tests of reading and vocabulary in the L2.

Findings. A comparison of global rating scales and detailed frequency scores as measures of specific aspects of oral L2 performance indicated that the two kinds of measurement were substantially correlated where there was sufficient variability in the data. An exploratory factor analysis of 26 variables, including measures of oral performance and academic test scores, yielded three orthogonal factors, interpreted as general English proficiency (including all the academic tests), vocabulary, and communicative style (consisting of interview variables). No separate factor was found for measures of fluency. Both the general English proficiency factor and the vocabulary factor were affected by length of residence in the L2 community, and general English proficiency was also affected by the students' age. Neither length of residence nor age was related to communicative style.

Conclusions. It was concluded that language proficiency results are strongly affected by the testing method (e.g. academic reading test, oral interview, story-telling task), and that an inherent difficulty in validating models of L2 proficiency is that measures faithfully reflecting a particular construct may not have adequate psychometric properties, while other psychometrically acceptable measures may fall short of representing the construct.

2:5 Metaphor Comprehension in Children's L1 and L2 (Final Report, Vol. I)

Purpose and design. This study compared the development of metaphor comprehension in Spanish-English bilingual children and monolingual English-speaking children, in order to test the hypothesis that metaphoric processing in bilinguals, as well as monolinguals, is constrained more by age and mental-attentional capacity than it is by language proficiency. Subjects were 20 Hispanic and 20 monolingual English-speaking children in each of three age groups: 7-8, 9-10, and 11-12 years, selected on the basis of a 'Figural Intersections Test' as being of normal mental capacity, which increased with age. An oral language proficiency test and a metaphor comprehension task in English were individually administered to each child. Hispanic children were also tested for oral proficiency in Spanish, and a subsample was tested for metaphor comprehension in Spanish. The language proficiency tests were similar to verbal IQ tests, while the metaphor comprehension task involved the oral interpretation of ambiguous metaphors, such as "my sister was a rock." The relative complexity of the children's metaphoric interpretations was coded with reference to the degree of semantic transformation involved in mapping an aspect of the vehicle (predicate) onto the topic (subject). The coding scheme had previously been shown to have good reliability and developmental validity for monolingual English-speaking children.

Findings. On the English proficiency test, Hispanic children scored significantly lower than the monolingual English-speaking children, and the Hispanic children resident in Canada for less than three years scored lower than those resident for more than three years. On the Spanish proficiency test, on the other hand, the more recent immigrants scored significantly higher than the long term residents. Performance on the metaphor comprehension task in English was, as predicted, found to be more strongly related to age and mental capacity scores than to oral language proficiency scores. While the bilingual Hispanic children did less well on the metaphor comprehension task than did the monolingual English-speaking sample as a whole, this was found to be related to the presence in the English-speaking sample of some students from a school in a higher socio-economic area. These children of middle-class background did better on the metaphor task than did the monolingual English-speaking children from the same schools as the bilingual children in working class areas. When the children of middle class background were removed from the sample, there was no main effect for language group on the metaphor scores, although the Hispanic children did less well on one of the two topics. Regression analyses indicated that the bilingual Hispanic children were similar to the subsample of English-speaking children from the middle-class neighbourhood in that English proficiency contributed little to the variance in their metaphor scores. Another finding was that conceptual structures developed in the first language appeared to facilitate metaphor comprehension in the second language, since for Hispanic children resident in Canada for less than three years, Spanish proficiency correlated more highly with metaphor scores in English than did English proficiency.

Conclusions. The findings of the study were in keeping with the hypothesis that, for bilingual as well as monolingual children, measured language proficiency was less predictive of metaphor performance than were age and non-verbal mental capacity scores. On a standardized test of English proficiency, the bilingual children scored significantly lower than their English-speaking schoolmates. On the metaphor task, however, the bilingual children performed almost as well as their English-speaking peers. This finding suggests that the metaphor task may be a more appropriate measure of conceptual skills in the second language than is a verbal IQ test.

3. CLASSROOM TREATMENT STUDIES

Several studies were undertaken to examine the relationship between instructional practices and the development of proficiency in a second language. During the first two years of the Project, a major focus was on the development and validation of a classroom observation instrument designed to capture the essential features of communication in the L2 classroom. This instrument was subsequently used in a process-product study which examined the impact on L2 proficiency of different instructional practices observed in core French classes. Two other studies grew out of the large-scale proficiency study described in 2:1 above. One of these involved the analysis of some specific aspects of language use and learning activities observed in French immersion classrooms, with a view to interpreting some of the earlier proficiency findings. The other study consisted of a classroom experiment in the French immersion setting, designed to enhance grammatical proficiency in the use of past tenses. These studies are summarized below.

3:1 Development and Validation of COLT Observation Instrument (Year 2 Report, Final Report, Vol. II)

The development of a new classroom observation scheme was motivated by the need to describe as precisely as possible some of the features of communication occurring in the second language classroom, and to distinguish between analytic and experiential orientations to language instruction. The COLT — Communicative Orientation of Language Teaching — scheme was derived from the communicative competence framework underlying the large-scale proficiency study and from a review of current issues in communicative language teaching.

Observation categories. The COLT observation scheme is divided into two parts. Part I, filled out by observers during the class, identifies different types of classroom activities and categorizes them in terms of: (a) participant organization (whole class activity, group work, individual work); (b) the content, or subject-matter, of the activity (e.g. classroom management, explicit focus on language form or function, other topics); (c) student modality (listening, speaking, reading, writing); and (d) materials in use (the type of material, length of text, intended users, and amount of control exerted on student language use). Part II of the COLT, which is later coded from a tape-recording of the class on a time-sampling basis, analyses communicative features of teacher-student interactions. Seven superordinate categories are identified: (1) use of target language (L1 or L2); (2) information gap (the level of predictability in an interaction); (3) sustained speech (length of utterances); (4) reaction to code or message; (5) incorporation of preceding utterances (how the participants react to each other's contributions); (6) discourse initiation (by teacher or student); and (7) relative restriction of linguistic form.

Validation. The observation scheme was piloted in 13 classes, mainly at the grade 7 level. There were four core French classes, two extended French and two French immersion classes, and five ESL classes in the sample. Each class was visited twice by two observers. Analysis of the Part I data entailed calculating the percentage of classroom time spent on the subcomponents of the various categories: participant organization, content, student modality, and materials. In the analysis of Part II, each verbal interaction feature was calculated as a proportion of its superordinate category. Results indicated that the COLT observation scheme was capable of capturing differences in the instructional orientation of the four types of classes. In core French

and ESL classes, for example, there was a heavier emphasis on form and more teacher control than in the extended French and immersion classes where there was a greater focus on meaning. Extended text was most often used in the immersion classes, and non-pedagogic materials were most often used in immersion and ESL classes. Sustained speech was least characteristic of the core French classes and most evident in French immersion and ESL classes. These comparative findings, intended as descriptive and not evaluative, generally met prior expectations concerning the various programs, except for some aspects of the ESL classes.

Conclusions. The ability of the COLT observation scheme to capture differences in instructional orientation was seen as an indication of its validity and as an important step toward identifying what makes one set of instructional techniques more effective than another.

3:2 The Core French Observation Study (Final Report, Vol. II)

In this process-product study, the COLT observation scheme was used to describe instructional practices in eight core French classes at the grade 11 level. Instructional differences were then analysed in relation to L2 proficiency outcomes in the different classes.

Subjects and procedures. The core French program was selected for study because the students' L2 proficiency could be assumed to derive largely from the classroom. The eight classes, from the metropolitan Toronto area, were preselected with the help of school board personnel to represent a range of L2 teaching practices. Early in their grade 11 year, the students were given a series of pre-tests of French proficiency, including some tasks from the large-scale proficiency study. The tests consisted of: (a) a multiple choice grammar test; (b) two written production tasks (a formal request letter and an informal note) which were scored for both discourse and sociolinguistic features; (c) a multiple choice listening comprehension test calling for the global comprehension of a series of recorded texts; and (d) an individual oral interview administered to a subsample of students from each class and scored for proficiency in grammar, discourse and sociolinguistics. During the school year, each class was visited four times for observation with the COLT scheme (in October, January, March and April). Observation periods lasted 40 or 70 minutes, depending on the duration of the class, and were tape-recorded. In May, the classes were post-tested with the same tests, and those students interviewed at the time of pre-testing were reinterviewed.

Analysis of COLT observations. Based on the Part I and Part II categories of the COLT observation scheme, it was possible to rank order the eight classes on a bi-polar composite scale from 'most experiential' to 'most analytic', based on the percentage of class time spent on practices defined as experiential in contrast to analytic. In the two most experiential classes, for example, there was proportionately significantly more topic control by students, more extended written text produced by the students, more sustained speech by students, more reaction (by both teacher and students) to message rather than code, more topic expansion by students, and more use of student-made materials than in the other classes. These two classes were labelled 'Type E' classes, in contrast to the remaining 'Type A' classes, where significantly more analytic features were in evidence, including a higher proportion of topic control by teachers, minimal written text by students, student utterances of minimal length, student reaction to code

rather than message, and restricted choice of linguistic items by students. The COLT analysis revealed at the same time that none of the classes was prototypically experiential or analytic, but instead intermediate along the bi-polar scale. The COLT findings were supported by teacher questionnaires providing information about classroom activities throughout the year.

The relationship of COLT findings to L2 proficiency. It was predicted that the Type A classes would be significantly higher on both written and oral grammatical accuracy measures than the Type E classes, but that the Type E classes would score higher on all other proficiency measures, including discourse and sociolinguistic measures, and scores on global listening comprehension. However, based on the post-test scores adjusted for differences in pre-test scores, no significant differences were found between the Type E and Type A classes, although a near-significant difference (p .06) emerged in favour of the Type A classes on the grammar multiple choice test. When the two Type E classes were compared to the two most analytic Type A classes (labelled Type A*), the Type A* classes did significantly better on the grammar multiple choice test (and specifically on agreement rules), but no other significant differences were found. A detailed correlational analysis relating the use of specific COLT features to L2 proficiency outcomes showed that there were both experiential and analytic features that were positively related to adjusted post-test scores. The profile of a successful core French classroom with respect to proficiency was identified as one in which a generally experiential approach was used with relatively more time devoted to features such as information gap, reaction to message, and topic incorporation. At the same time, positive correlations were found between a number of form-focussed, teacher-directed activities and proficiency outcomes. It was concluded that analytic and experiential approaches may be complementary rather than two ends of a continuum.

Qualitative analysis. In light of the fact that one of the two Type E classes made the most gain in overall proficiency and that the other Type E class made the least gain, the transcripts of these two classes were reviewed for qualitative differences that had not been captured by the COLT. The high-scoring class was found to engage frequently in communicatively rich interaction involving feedback and negotiation of meaning, while the low-scoring class received less feedback and spent more time on stereotyped routines. It was concluded on the basis of these findings that observation procedures based on COLT needed to be supplemented by more detailed forms of discourse analysis.

3:3 The Immersion Observation Study

(Final Report, Vol. II)

Classroom observations were carried out in nine grade 3 and ten grade 6 early total immersion classes in the Toronto and Ottawa regions, for the purpose of obtaining information on classroom processes and interpreting earlier immersion L2 proficiency findings. Each class was observed and tape-recorded for a full school day, including any instruction in English. All the tape-recordings were then transcribed. Analyses of some aspects of language use in the immersion classes are presented in the Project report. Further analyses are planned as time and finances permit.

Vocabulary instruction in immersion classes. L2 vocabulary-related classroom activities observed in the grade 6 immersion classes were analysed in the light of a classification scheme for describing such activities, and in relation to different kinds of linguistic knowledge involved in vocabulary learning. The analysis is qualitative rather

than quantitative. The classification scheme focusses on the following distinctions: (a) planned/unplanned instruction — the extent to which vocabulary instruction is seen as a purposeful activity; (b) systematic/haphazard instruction — the degree of systematicity with which specific features of vocabulary are taught; (c) written/oral activities — use of each medium for vocabulary instruction is seen to have a differential effect on lexical knowledge; (d) cross-linguistic/L2 based approaches to vocabulary teaching — a role for controlled use of the L1 is noted; (e) control of vocabulary selection — this may be by text authors, teacher or students; (f) the linguistic focus of teaching — whether the focus is on interpretation in context, conventional meaning, or other structural aspects of vocabulary. Linguistic aspects of vocabulary knowledge are categorized in terms of phonology, morphology, syntax, discourse semantics, graphology, and sociolinguistics. Analysis of the immersion classes in the light of these descriptors indicated that most planned vocabulary teaching occurred during reading activities organized around particular themes, during which students learned to pronounce words that they read aloud, to interpret passages, and in which the meanings of unfamiliar words were explained. Unplanned, spontaneous teaching of vocabulary was often student-initiated, as a specific word was needed to express an idea. The focus of both the planned and unplanned vocabulary teaching was mainly on interpretation and meaning. Given its association with reading activities, the teaching of new words emphasized written varieties of French and spelling. One example of how the students' prior knowledge could be drawn upon was provided by a teacher who drew attention to formal resemblances in the L2, enabling students to arrive at the meaning of an unfamiliar derived word. With some exceptions, the presentation of structural information about vocabulary was limited to the separate grammar lesson. Errors in gender, for example, tended to be only haphazardly corrected during other activities. There was no evidence that the teachers were focussing on sociolinguistic or discourse-related aspects of vocabulary. It was concluded that vocabulary teaching in the immersion classes occupied a rather narrow place in the overall teaching plan, and was mainly preoccupied with meaning and interpretation with insufficient planned attention to other aspects of vocabulary knowledge.

Vous/tu input. The underuse of vous as a politeness marker by early immersion students had been noted in the earlier proficiency study. The classroom observations provided an opportunity to relate these findings to use of vous and tu in the classroom context. Uses of tu and vous by the ten grade 6 immersion teachers and by the students in their public talk were counted and classified according to the functions they served: singular, plural, or generic; formal or informal. Teachers were found to use tu and vous about equally often, with tu generally being used to address individual children and vous to address the class as a whole. Occasionally, however, tu was used to the class and vous to individual students, leaving room for potential confusion. There was scarcely any use of vous by the teachers as a politeness marker, and its infrequency in this function in the classroom context was seen as a reason for its underuse as a politeness marker by early immersion students. Their underuse of vous in its plural function was at the same time attributed in part to the finding that, although used relatively frequently by teachers, very few opportunities appeared to arise for student production of vous plural in the classroom context. In conclusion, it was hypothesized that students would benefit from functionally-oriented instruction in the use of vous/tu, and opportunities to use them appropriately.

Student talk in teacher-fronted activities. It had previously been hypothesized that shortcomings in the grammatical competence of early immersion students may be due to

a lack of opportunity to produce 'comprehensible output', i.e. precisely conveyed messages demanding more rigorous syntactic processing than that involved in comprehension. In order to determine the opportunities that the immersion students had to talk in class, transcripts based on 90 minutes of French class time in each of the nine grade 3 and ten grade 6 classes were analysed, as well as the English portion of the day in the grade 6 classes. In general, the transcripts captured public talk rather than private, individual conversations. Each student turn was categorized according to length (minimal, phrase, clause, and sustained), and source (e.g., whether teacher- or student-initiated, preplanned or unplanned, linguistically restricted in some way or not, reading aloud from text, or reciting from memory). The findings indicated that in the 90-minute French portion of the day, student talk was less than two thirds as frequent as in the English portion of the day. Sources of student talk in French were very similar for the grade 3 and grade 6 students, the most frequent source being teacher-initiated student talk where the students' response was highly linguistically constrained, which appeared to encourage minimal responses from the students. Extended talk of a clause or more appeared to be encouraged when students initiated talk and when they had to find their own words. However, less than 15% of student turns in French were found to be sustained, i.e. more than a clause in length, when reading aloud was not included. It was concluded that greater opportunities for sustained talk in French by the immersion students are needed, and that this might be accomplished through group work, the provision of more opportunities for student-initiated talk, and through the asking of more open-ended questions by teachers.

Error treatment. An analysis of the grade 6 immersion teachers' correction of errors was based on the complete French transcripts of the ten classes observed. It focussed on the grammatical and pronunciation errors corrected by the teachers, the proportion of such errors corrected, and the systematicity of error correction. The highest proportion of error was observed in frequently used grammatical features such as gender, articles, and verbs. Only 19% of grammatical errors overall were corrected, but gender, article, and verb errors were more often corrected than other grammatical errors. About two-thirds of pronunciation errors were corrected. A lack of consistent and unambiguous teacher feedback was noted.

3:4 Functional Grammar in French Immersion (Final Report, Vol. III)

This experimental study was designed to investigate the effect on immersion students' French proficiency of an approach to grammar teaching which involved the provision of focussed input in a problematic area of French grammar and provided students with increased opportunities for meaningful productive use of the target forms. Following a workshop with teachers, a set of classroom materials aimed at teaching the meaning distinctions between two major past tenses, the imparfait and the passé composé, were introduced for an eight-week period into grade 6 early immersion classes in six schools. These experimental classes were compared on pre-tests, immediate post-tests, and on delayed post-tests (three months later) with comparison grade 6 immersion classes in six other schools who were not exposed to the materials. The tests consisted of narrative compositions previously used in the large-scale proficiency study, as well as specially constructed cloze tests with rational deletions, and oral interviews administered to a sub-sample of students in each class. All the tests were designed to assess the students' ability to make appropriate use of past tenses and were scored

accordingly. There were two forms for each test, administered across testing sessions in a counterbalanced design.

The classroom materials. Adapted from an existing bank of activities focussing on the imparfait and the passé composé, the materials were divided into eight units, each to be used in a specific week. The teaching approach emphasized the integration of grammar teaching with worthwhile subject matter content and the personal experience of students. The oral and written activities, providing focussed input and opportunities for practice in using the two tenses, included the following: reading a simplified French-Canadian legend, discovering how the imparfait and passé composé served different functions in the legend, illustrating aspectually contrasting sentences, applying proverbs to the legend and to the students' own experiences, miming the progressive function of the imparfait, working in small groups to create new legends, and producing albums of childhood memories.

Findings. On the immediate post-tests, with adjustment made for pre-test scores, the experimental classes were significantly ahead of the comparison classes on two out of three measures: the cloze test and the oral interview. Three months later, however, at the time of delayed post-testing, there were no significant differences between the experimental and comparison groups on any of the tests. Both groups had improved their test performance over time. Evaluations of the materials by the experimental teachers at the end of eight weeks indicated general satisfaction with the materials, although some problems were noted with specific activities. Teachers indicated that they spent on average about 1 1/2 hours per week on the material. From some of their comments, it appeared that certain activities promoted more attention to subject matter content than to linguistic code, and informal observations in some classes indicated that past tense errors often went uncorrected during the 'Proverbes' activity. It was noted that one class with a teacher who was observed to provide frequent corrective feedback obtained the best results of all the classes on the composition test. Questionnaires administered to experimental and comparison group teachers at the time of the delayed post-testing indicated that the latter had also spent time working on the target verb tenses.

Conclusions. It was concluded that the teaching approach had succeeded in accelerating grammatical development in the experimental classes, but that to promote more long-term benefits some revision was needed in the materials, including more specific guidelines to teachers about the provision of corrective feedback. The fact that the control classes also appeared to have worked on past tenses was an additional factor that was surmised to have affected the long-term results.

4. SOCIAL CONTEXT AND AGE

The relationship between individual and social-environmental factors and the development of bilingual proficiency was examined in several minority and majority language learning contexts. In one large-scale study of Portuguese-Canadian students, the relationship between language use patterns, language attitudes, and bilingual proficiency was investigated by means of correlational and regression analyses, while in a small sample of beginning school-aged children of Portuguese home background, a detailed study of language interaction at home and at school was carried out with a view to relating interactional variables to later academic achievement. In another minority context, an ethnographic study focussed on students attending a French language

elementary school in Toronto. Finally, two studies examined the relationship between age and language learning: one among Japanese immigrant students of different ages and the other among Anglophone majority students learning French in three different school programs.

4:1 Language Use, Attitudes and Bilingual Proficiency of Portuguese Canadian Children (Final Report, Vol. III)

Purpose and design. In this study, the bilingual proficiency of grade 7 students from an important language minority group in Toronto was studied in relation to family background variables, the students' patterns of language use, and their language attitudes. Theoretical issues examined were: (a) the nature of language proficiency indicated by the pattern of relationships within languages; (b) the cross-lingual dimensions of language proficiency indicated by the pattern of relationships across languages; and (c) the extent to which proficiency in English and Portuguese could be predicted by language use and attitude variables.

The sample consisted of 191 students enrolled in Portuguese heritage language programs in seven inner-city Toronto schools. More than half these students were of Azorean background. The students all completed two questionnaires. One was a language use questionnaire concerning family background (e.g. birthplace, parents' language use, education, and occupations), language use patterns (use of Portuguese and English at home, in school, and in the community), and self-ratings of proficiency in English, Portuguese, and French. The other was a language attitude questionnaire which investigated dimensions such as integrative and instrumental orientations towards English and Portuguese, language use preferences in different contexts, the role of English and Portuguese in the students' ethnic identity, perceived attitudes of parents towards the students' education and language use, attitudes towards Portuguese dialects and language mixing, cultural assimilation, and attitudes towards French. Tests in English and Portuguese were also administered. In each school the students were divided randomly into three groups. One group did multiple choice grammar tests in English and Portuguese. A second group received a multiple choice discourse test in each language similar to the one administered in the large-scale proficiency study (see 2:1 above). Students in this group were also given individual oral tests in English and Portuguese, each of which contained tasks to be scored for grammar, discourse, and sociolinguistic proficiency. The sociolinguistic task in each language was adapted from the oral sociolinguistic test administered in the large-scale proficiency study. A third group of students in each school was given sociolinguistic written production tests in each language, again based on the test designed for the large-scale proficiency study.

The nature of language proficiency and its cross-lingual dimensions. A considerable degree of interrelationship was found among Portuguese self-ratings, multiple-choice discourse scores in Portuguese, and the various oral measures of Portuguese proficiency. A principal components analysis suggested a global Portuguese proficiency dimension, supplemented by academically related aspects of proficiency. Few relationships, on the other hand, were found among the measures of oral English proficiency, apparently because of a generally high level of performance giving rise to a lack of variability in scores. Across languages, self-ratings of proficiency in Portuguese, English, and French tended to be significantly related to each other. Further relatively strong cross-lingual relationships were observed for each set of written measures: i.e. between multiple choice grammar scores in English and Portuguese, between multiple

choice discourse scores in each language, and between written sociolinguistic scores in each language. These findings provided strong evidence for the interdependence of cognitive-academic skills across languages.

Predictors of bilingual proficiency development. Multiple regressions revealed that a considerable amount of the variance in the self-ratings of Portuguese proficiency could be related to attitudinal and language use variables such as students' acceptance of Portuguese, their knowledge and pride in Portuguese culture and achievements, their use of Portuguese media, exposure to Portuguese in the home, and their acceptance of and liking for French. Ceiling effects on the English self-ratings appeared to be at least partly responsible for the much weaker correlations found with attitude and use variables, although positive relationships were found with acceptance of English, use of English with siblings, and acceptance of French. The amount of exposure to Portuguese, both in Portuguese language classes and in the form of visits to Portugal, attendance at Portuguese mass, and Portuguese TV watching, appeared to be strongly related to measures of Portuguese proficiency, with weaker relationships noted between attitude variables and Portuguese proficiency. Minimal relationships were found between language use and attitude variables and the English proficiency measures, although there was evidence to suggest that positive attitudes towards Portuguese and students' use of Portuguese at home and in the community were in no way detrimental to their English proficiency.

Comparison with Azorean native speakers. A comparison of the Toronto students' test scores in Portuguese with those obtained by 69 grade 6 students in the Azores revealed that there were highly significant differences favouring the Azorean group on most measures of Portuguese proficiency. As in the large-scale proficiency study involving French immersion students (see 2:1 above), differences were most apparent on measures of grammar. The strong relationship found between Toronto students' attendance at Portuguese language classes and proficiency in Portuguese was seen as an indication that, in their minority context, more intensive exposure to Portuguese in an academic context could be advantageous for the bilingual development of the Toronto students.

4:2 Longitudinal Study of Young Portuguese Background Children: Bilingual Proficiency Development and Academic Achievement (Final Report, Vol. III)

Purpose and design. The major purpose of this ongoing study is to investigate the development of proficiency in both Portuguese and English in the transition from home to school. Twenty children from Portuguese backgrounds are being followed from the junior kindergarten year through grade 1 with respect to patterns of language interaction in the home, performance on a variety of language proficiency and literacy awareness measures, and (in grade 1) reading performance. Patterns of interaction in the home and knowledge of Portuguese and English will be used as predictors of English reading performance in grade 1. Thus, the study addresses theoretical issues such as the interdependence of L1 and L2 as well as practical issues related to the interaction between home and school variables in affecting the extent to which minority students are successful academically. The study also will provide a corpus of longitudinal data for analysis of students' developing proficiency in their two languages.

Methodology. The main sample consists of 20 Toronto students receiving the entire battery of tests. These are the Draw a Person Test, the Record of Oral Language (i.e.

sentence repetition) (English and Portuguese), Letter Identification (English and Portuguese), Concepts about Print (English and Portuguese) and, in Year 3 (Spring 1987), Test of Writing Vocabulary (English and Portuguese) and Gates McGinitie Reading Test (Comprehension subtest). (For test references, see complete study in Chapter 8.) In addition, children were taped in their homes for one and a half hours each year of the study.

Twenty-six grade 1 students (average age 7) in the Azores were also administered the Concepts about Print test, an oral interview, and Test of Writing Vocabulary (Clay 1979) in Portuguese for comparison purposes with the grade 1 Toronto data. In addition, six five-year-old students in the Azores were taped for one and a half hours in their homes. Data were also collected in Mainland Portugal from ten five-year-old children in a village situated a hundred kilometres northwest of Lisbon. A Portuguese version of the Record of Oral Language was constructed and administered to the children. Six of the ten were randomly chosen to be taped in the home.

Current status of the study. All the Year 1 home recordings have been transcribed and scoring schemes developed for grammar and pragmatics. A sample of students' transcripts have been scored but not the entire group. The Year 3 data will be collected in May and June of 1987. Subsequent to this data collection, a proposal will be developed to complete the transcription and data analysis relating home language use and proficiency in literacy-related aspects of English and Portuguese to English reading performance at the grade 1 level.

4:3 Ethnographic Study of a Toronto French Language School (Year 2 Report; see also Heller 1984)

In this ethnographic, sociolinguistic case study of a French-language elementary school in Toronto, patterns of language choice and language use were investigated in relation both to the micro-level interactional context and to the macro-level context of school and community. The study examined the role that the use of French and English played in the development of students' social identities.

Methodology. Micro-level data were collected in the school by means of participant observation over a six-month period, mainly in a grade 7/8 class, and through tape-recordings of eight students who each wore a tape-recorder for two entire school days. Four of the students were selected as ethnolinguistically representative of the school and the other four were randomly selected. Macro-level data were collected through a school-wide parent questionnaire and in interviews with school administrators, staff, members of the Parent-Teacher Association, and an ethno-linguistically representative subsample of parents.

Findings. Just over half the parents returned their questionnaires, which indicated considerable heterogeneity of family origins, linguistic backgrounds, and goals with respect to bilingualism and the maintenance of French. For example, over 40% of the families were of linguistically mixed marriages (usually with a francophone mother), 30% were francophone, 11% anglophone, and the remainder from a great variety of linguistic backgrounds. Very few parents and under half the children were Toronto-born. Family homes were widely dispersed over half of the city, making it hard for students to maintain friendships outside school. In-school observations revealed that there were three distinct groups of students: English-dominant, bilingual, and French-dominant.

The first two preferred to speak English among themselves, and the third -- a minority -- preferred French. Access to the different peer networks depended on appropriate language choice. Each group experienced its own tensions: French-dominant students reported pressure from peers to speak English outside class, while for English-dominant students, performance in French in class could be stressful. Bilingual students were observed to take part in occasional bilingual word-play and code switching, which was seen as their way of resolving the social tensions they experienced from their intermediate position and suggested that, for them, French and English were separate domains.

Conclusions. The heterogeneity of the school population and the varied linguistic experiences of the students were seen to militate against the formation of a monolithic French identity. Instead, observed patterns of language use indicated a close connection for the students between language choice and their evolving social identities.

4:4 Age on Arrival, Length of Residence, and Interdependence of Literacy Skills among Japanese Immigrant Students (Final Report, Vol. III)

Purpose and design. This study investigated the cross-lingual dimensions of language proficiency and the relationship between age and second language acquisition, with a focus on the development of reading and writing skills. We hypothesized that despite the dissimilarity of languages and writing systems, significant positive relationships would be found between Japanese minority children's L1 reading and writing skills and their acquisition of English reading and writing. An investigation of the relationships between Japanese and English proficiency appears to provide a stringent test of the interdependence hypothesis, which posits a common underlying proficiency for bilinguals, since the two languages have little in common at a surface structure level.

Subjects in the study consisted of 273 students between grades 2 and 8 attending the Japanese School of Toronto Shokokai Inc. Students were tested in May and June 1984 with measures of reading and writing in both Japanese and English. The reading comprehension subtest appropriate to students' grade level of the Gates-McGinitie Reading Test was given to all students who had been in Canada for at least six months as a measure of English reading skills. The Kyoken Standardized Diagnostic Test of Reading Comprehension published by the Research Institute for Applied Education in 1981 was given as the measure of Japanese reading skills. In addition, a letter-writing task in English and Japanese was administered to all children.

Scores on the English and Japanese reading tests were converted to T-scores to permit comparability across grades with the influence of age removed. In addition, English grade equivalent scores were used in some analyses as an approximate index of students' absolute level of English reading skills. A variety of indices of writing skills in Japanese and English were assessed.

Results. The results of correlational and regression analyses provide a consistent picture in relation to the acquisition of English reading and writing skills and their relationship to students' Japanese reading and writing proficiency. First, although the sample as a whole performs close to the mean (i.e. Japanese norms) in Japanese reading skills, there is a clear negative relationship between length of time in Canada and students' Japanese reading proficiency. The negative effect of length of residence on Japanese writing, however, appears minimal. Age of arrival in Canada appears to be a

more potent force in predicting maintenance of Japanese writing skills than length of residence. Similarly for Japanese reading, the older students are when they come to Canada, the better prospects they have for strong continued development of Japanese reading skills. This effect is not entirely due to the fact that students who arrive at older ages tend to have spent less time away from Japan, since the partial correlation between age of arrival and Japanese T-score remains significant even when length of residence is controlled.

It appears that students require about four years' length of residence, on the average, to attain grade norms in English reading skills. There appears to be some tendency for students who arrive at the age of 6-7 to make somewhat more rapid progress towards grade norms than those who arrive at older ages.

When length of residence is controlled, a significant relationship emerges between Japanese reading skills and English reading. Students' age of arrival in Canada (AOA) is also strongly related to English reading (controlling for length of residence), suggesting the influence of general cognitive maturity in mediating the cross-lingual relationship of cognitive/academic skills. General cognitive maturity, however, cannot account fully for the interdependence of reading skills across languages since significant relationships across languages were found for reading T-scores, in which the effects of age have been removed.

Writing performance was less closely related across languages than was the case for reading. This may be partly a function of the different types of measures used in each case (standardized reading tests v. non-standardized writing tasks). However, consistent significant relationships were obtained between Japanese writing and both English reading and writing measures. For some variables (e.g. Spelling) there was strong evidence of a specific cross-lingual relationship that was not mediated by more general cognitive/academic proficiencies.

Conclusions. In general, the data are consistent with previous studies in supporting the interdependence of cognitive/academic skills across languages. They also suggest that at least four years is required for students from highly educated backgrounds to attain grade norms on English academic tasks and that continued development of L1 academic skills to a high level (i.e. that of students in the home country) is a formidable task for students who arrive in the host country at an early age (particularly prior to formal schooling) but is considerably less problematic for students who arrive after several years of schooling in their home country.

4:5 Starting Age and Oral French L2 Proficiency in Three Groups of Classroom Learners (Final Report, Vol. III)

Purpose and design. The purpose of this study was to determine whether there are specific long-term advantages in oral L2 proficiency that can be associated with intensive L2 exposure at an early age in a total French immersion classroom setting. Three groups of grade 10 learners, with 11-12 subjects per group, were interviewed and given an oral sociolinguistic test in French: one group was from an early total immersion program which had begun in kindergarten, while the other two groups (from a late immersion and an extended French program respectively) had started their intensive exposure to French much later, in grade 7. A group of 12 native French speakers in grade 10 was also included in the study. The guided oral interviews were designed to

provide students with communicative contexts for the use of a range of verbs and verb forms. Transcripts of the interviews were analysed with respect to verb use and oral fluency in French. Scoring of verbs consisted of assessing the use of target verb forms in the context of specific questions, while the assessment of oral fluency was based on the nature and frequency of markers of disfluency and the linguistic contexts in which they occurred. The sociolinguistic oral test was based on the one used in the large-scale proficiency study (see 2:1 above).

Results. Group comparisons of the students' verb use indicated that the early immersion students were significantly more native-like on some variables (imparfait, conditional, use of pronoun complements in clitic position), but were no more native-like than the other learner groups on other variables such as use of number and person distinctions, time distinctions, and lexical variety, and in some instances tended to be less native-like than one or both of the other groups. The analyses of fluency revealed that in most types of disfluency, the three learner groups produced significantly more disfluencies than the native speakers but did not differ from one another. There was some evidence, however, that the early immersion students were producing fewer cut-offs and 'uh', 'um' etc. transition markers. The early immersion students were also less likely than the late immersion students to use transition markers in within-phrase locations, where such disfluencies were hypothesized to be more disruptive to discourse coherence than in between-clause or between-phrase locations. These findings indicated some advantages in oral fluency for the early immersion students who had started their intensive L2 program at a young age. Results on the sociolinguistic oral test, however, showed that the early immersion students did not manifest any general advantage over the other learner groups in sociolinguistic proficiency. While the early immersion groups displayed a slightly greater tendency to use attenuating conditional verb forms in formal social situations, they tended to be less sensitive to the appropriate use of the second person forms vous and tu than the late immersion and extended French students, whose intensive exposure to French in school had begun much later.

Conclusions. With respect to oral L2 proficiency, it appeared that there were some advantages to an early start in a French immersion program in the area of fluency and in the use of the verb system, but no advantage in the sociolinguistic domain. Some weaknesses in the verb system were also observed. As in other studies conducted in the early immersion context, a need for more emphasis on problematic areas in the target language system was indicated, along with greater opportunities for sustained oral and written expression.

5. CONCLUSIONS

The overall conclusions of the studies can be summarized with respect to the nature of bilingual proficiency and the influences on its development both in classroom and natural settings.

The nature of proficiency. At the inception of the study, the primary methodology envisaged for investigating the nature of language proficiency and its cross-lingual dimensions was confirmatory factor analysis. However, as a result of the findings of our Years 1-2 study of proficiency among French immersion students, in which little evidence emerged for the hypothesized trait structure, we became more explicitly conscious of the fact that the relationships between different components of language proficiency were a function of the specific language learning experiences to which

particular samples of individuals were exposed. This perspective implies a wider variety of analytic methods for investigating the nature of proficiency; specifically, we can discover a considerable amount about the nature of proficiency by observing its behaviour as a function of individual, social and educational conditions. Thus, we shall first consider the findings of our factor analytic studies and then examine findings of other studies that elucidate the nature of proficiency.

All studies that examined the relationships among different components of proficiency found significant correlations among written tests (including the core French observation study -- see 3:2 above). These relationships were found across languages in the grade 7 Portuguese study (4:1), the Japanese study (4:4), and the metaphor comprehension study (2:5). Some evidence emerged for an oral factor (e.g. a communicative style dimension in the "Communicative skills of young L2 learners" study -- 2:4) but the relationships among oral measures were considerably less strong than for the written measures. Similarly, some cross-lingual relationships among oral measures were found in the Portuguese grade 7 study but again the relationships were only marginally significant. These data are consistent both with the notion of a specific dimension of proficiency related to the ability to process language in context-reduced or decontextualized situations and with the hypothesis that this dimension is interdependent across languages.

There was considerably less evidence in the factor analyses for the hypothesized trait structure distinguishing grammatical, discourse, and sociolinguistic aspects of proficiency. However, discourse and grammar factors did emerge in the lexical proficiency study (2:3) and there was also some evidence for a separate vocabulary factor. The limitations of placing exclusive reliance on factor analysis for confirming hypothesized trait structures are illustrated in the fact that in this lexical study several mutually exclusive solutions produced an acceptable fit to the data. Also, in the original proficiency study (2:1), comparison of French immersion with native French speakers produced evidence that discourse skills were distinguishable from grammar and sociolinguistic skills, in that differences between L2 learners and native speakers were found only for the latter two aspects of proficiency.

Thus, consistent with the position advanced by Cziko (1983), the lack of strong support for the hypothesized trait structure in the factor analyses does not lead us to abandon the concept of traits. They are conceptually distinguishable and educationally important even if they are not statistically verifiable in relatively homogeneous school populations.

Classroom treatment. Our classroom treatment findings from different program settings lead to three main overall conclusions. First, there is evidence from both the core French and the immersion observation studies that the analytic focus and the experiential focus may be complementary rather than two ends of a continuum, and that they may provide essential support for one another in the L2 classroom. Second, the quality of instruction is clearly important in both analytic and experiential teaching. Analytic teaching will be successful in developing L2 proficiency only if it is appropriately matched to the learners' needs, while experiential teaching should involve communicatively rich interaction which offers plenty of opportunities for production as well as global comprehension on the part of the student. Third, learners may benefit if form and function are more closely linked instructionally. There is no doubt that students need to be given greater opportunities to use the target language.

Opportunities alone, however, are not sufficient. Students need to be motivated to use language accurately, appropriately, and coherently. In all these respects, the 'how' and 'when' of error correction will be a major issue for future investigation.

It seems reasonable to conclude that in all the programs under investigation — core French, heritage languages, and French immersion — much more work needs to be done in the area of curriculum design. Such work should include research to determine what combinations of analytic and experiential activities are most effective for different types of student. Another comparatively neglected area from the research point of view is teacher training and professional development. This area is likely to become more important at a time when more and more teachers are breaking away from their former dependence on prescribed pedagogic formulas and are increasingly making their own, more flexible, decisions about what can be done in the classroom.

Individual and social variables. With respect to the influence of individual and social variables on the development of proficiency, we can think of these effects in terms of the relative influence of attributes of the individual (e.g. cognition, personality) versus the target language input received by the individual. With respect to attributes, for example, it is clear from the Portuguese grade 7 and Japanese studies (4:1 and 4:4 above), as well as the immersion age study (4:5) that cognitive attributes of the learner play a significant role in at least certain aspects of target language acquisition. In the grade 7 Portuguese study and the Japanese study, children's cognitive/academic proficiency in their L1 was significantly related to the level of cognitive/academic proficiency attained in the L2. The relatively strong performance of late immersion students in comparison to those in early immersion is consistent with the notion that the learner's cognitive maturity (as indicated by age) is positively related to efficiency of L2 acquisition (at least up to the point where cognitive development reaches a plateau, possibly in the early to middle teens).

There is some evidence that cognitive attributes are more related to acquisition of certain aspects of proficiency than to others. For example, L1 cognitive/academic skills are more closely related in the Portuguese grade 7 study to performance on L2 written (context-reduced) tasks than is the case for oral tasks. Also, discourse proficiency appears to be somewhat less influenced by input/exposure variables than is the case for grammar, as illustrated by the native-speaker comparisons in the large-scale proficiency study (2:1) and Portuguese grade 7 study as well as in the regression analyses for Portuguese proficiency in the latter study (4:1).

In short, one way of thinking about the trait structure and its relationship to psychological variables is to distinguish between aspects of proficiency that are relatively more dependent on input from the environment for their full development than on attributes of the individual (e.g. oral grammar) and those that rely probably as much on individual attributes (e.g. cognitive skills, personality variables) as on input for their development (e.g. oral and written discourse, context-reduced proficiency generally). We would see sociolinguistic aspects of proficiency (particularly in the oral mode) as intermediate between grammar and discourse with respect to their relative dependence on input versus attributes. In the case of sociolinguistic proficiency, personality variables are likely to be at least as important as cognitive variables but input is clearly also crucial, as demonstrated by the immersion observation study (3:3), which showed minimal input to students regarding sociolinguistic variation. The relatively greater problems that early immersion students experience with grammar and sociolinguistic

proficiencies in comparison to discourse is consistent with this position, as is the more evident influence of exposure variables (e.g. visits to Portugal) on grammar than on discourse skills in the Portuguese grade 7 study.

In conclusion, the picture of bilingual proficiency that emerges from our studies is one of a dynamic evolving complex of traits that become differentiated from each other as a function both of variation in the input from the classroom or wider environment and the individual attributes of the learner.

Footnote

[1] In recognition that abstract, underlying language competence is not directly measurable, but inevitably coloured by the method of elicitation used, the term 'proficiency' is used in this report in a global sense to encompass both competence and performance aspects of grammar, discourse, and sociolinguistics that are measured by our tests.

References

Cziko, G. Some problems with empirically-based models of communicative competence. **Applied Linguistics**, 1983, 5, 23-38.

Heller, M. Language and ethnic identity in a Toronto French language school. **Canadian Ethnic Studies**, 1984, 16(2), 1-14.

Chapter 1

PERSPECTIVES ON LEXICAL PROFICIENCY IN A SECOND LANGUAGE

Birgit Harley, Mary Lou King and Jud Burtis

1. INTRODUCTION

There is abundant evidence that lexical proficiency, or knowledge of 'lexemes' (Lyons 1977: 19), is a crucial aspect of communicative competence in a second language (L2). One kind of evidence comes from studies of native-speaker reactions to learners' errors: for example, in three separate studies conducted by Johansson (1978), Lepicq (1980) and Politzer (1978), an identical finding was that native speakers rated lexical errors more severely than errors of other kinds. Each of these studies included several categories of morpho-syntactic as well as lexical errors, and in Politzer's study, phonological errors were also among those rated. From the comments of raters, Lepicq concluded that comprehensibility of a learner's message was a major factor in the native speaker evaluations.

Another indication of the importance of lexical proficiency to communication in a second language comes from observations of learners' behaviour in conversation. Studies of communication strategies (see e.g. Faerch and Kasper, 1983) show that learners perceive the problems they have in making themselves understood as primarily lexical in nature. Learners' appeals for assistance, gestures, and paraphrases, for example, focus on gaps in vocabulary rather than on grammatical structure or pronunciation difficulties, as is evident in the following examples from French immersion students in grade 5 (Harley and Swain, 1978):

(1) j'ai um ... 'spray paint'?

(2) le devant de le car (makes crashing noise)

(3) et le bateau a — euh um était comme un sousmarin (meaning that the boat 'capsized').

The focus of learners on content words rather than grammatical functors may not only be observable from their speech behaviour, but can also emerge as a learning/communication strategy explicitly stated by a learner in an untutored learning environment. Shapira (1978) reports on one such learner in the U.S.A. who says: "I'm hear and put more attention the big words," and that "little words is no too important for me." The learner in question was found to have made no progress in acquiring grammatical morphemes in English over a four-month period of observation. (The development of her vocabulary was not, however, assessed.)

Such empirical observations raise interesting questions about the relationship of lexical proficiency to other aspects of L2 communicative competence. Is a high degree of lexical proficiency necessarily accompanied by superior control of morpho-syntactic structures, or does one develop relatively independently of the other? What relationship does lexical proficiency have to other aspects of communicative competence such as discourse competence (the ability to interpret and produce discourse coherently and cohesively) or sociolinguistic competence (the ability to interpret and produce language that is socially appropriate in context) (Canale 1983; Year 2 Report).

The present study was designed to examine these and other related issues concerning the development of lexical proficiency in French by English-speaking students enrolled in an early French immersion program.[1] The purpose of the study, which focusses on the verb vocabulary used in written compositions, is threefold:

(1) to compare different quantitative measures of the students' lexical proficiency in their second language;

(2) to examine the relationship between lexical proficiency and other aspects of their L2 communicative competence; and

(3) to describe the learners' lexical use in relation to that of native speakers of French in order to gain insights into the L2 lexical acquisition process.

2. THE DATA BASE

As part of a study investigating the factorial structure of second language proficiency in a sample of 69 grade 6 immersion students in six different classes (Swain 1985, Year 2 Report), the students were required to produce several written compositions. These compositions provide a substantial set of data from each student which can be further examined from a lexical perspective. In addition, comparison data are available from a grade 6 class of native speakers of French in Montreal. The time assigned for each of three composition-writing sessions was 30 minutes, with 15 minutes allotted to a narrative and 15 minutes to a request letter in each of two testing sessions, and 15 minutes to a request letter and 15 minutes to two short notes in a third session. These writing tasks were originally designed to be scored for grammatical, discourse, and sociolinguistic competence as indicated in Figure 1, p. 60. The instructions given for each task supplied the students with much of the noun vocabulary they needed to write their compositions, while the notes offered little opportunity overall to demonstrate lexical proficiency. For the purposes of the present study of lexical proficiency, therefore, it was decided to focus on the use of verb vocabulary in the two narratives and three request letters, which provided ample opportunities to demonstrate proficiency in the use of this major open lexical class.

Topics for the narratives were (1) the rescue of a kitten from a tree, and (2) a bank robbery. The letters, all addressed to adults, were on the following topics: (1) a request to keep a puppy in an apartment; (2) a request to borrow a bicycle; and (3) a request for information on pollution for a school project. The exact instructions given to students for the two narratives and three letters are reproduced in Appendix A, p. 64.

As a preliminary step in the analysis of the use of verb lexis, a list of items was compiled for each student, registering each occasion of use ('token') of a particular verb lexeme, or verb 'type', as well as each lexical error that was made, and identifying the composition in which it occurred. Any instance of a verb lexeme in finite, infinitive, or participle form was included in the list; thus il est allé, nous allons, aller, en allant all counted as tokens of the lexical verb type, aller. Participles functioning as adjectives were also included: for instance, fatigué in l'homme fatigué. When used as a predicate, constructions such as l'homme est fatigué, il est mort were identified as finite instances of the verb types fatiguer and mourir respectively.[2] Not counted as tokens were forms of être, avoir and aller when used as auxiliaries, and frozen verb forms occurring as part of a fixed expression (e.g., s'il vous plaît, est-ce que). The occurrence of each verb token was listed in the infinitive form of the relevant verb lexeme. Each item was at the same time coded as to whether it had occurred in an obligatory context for a finite verb, an infinitive, or a participle.

In general, **le français fondamental** (Gougenheim, Michéa, Rivenc and Sauvageot 1964) was used as a guide in determining what constituted a distinct verb type. This basic French word frequency list contains more than 1,000 word types based on an oral corpus, of which roughly 200 are verbs. The occasional distinction that this frequency list makes between reflexive (pronominal) and non-reflexive verbs was dropped for the purposes of the quantitative analyses. The reason for this was twofold:

(1) immersion students often did not make a distinction between reflexives and non-reflexives (e.g. amuser might be used instead of s'amuser).

(2) errors in pronouns and auxiliaries of reflexive verbs had been treated as grammatical errors in Allen et al. (1983) and therefore were not scored as lexical errors. It would not have made sense to give a student credit for having two correct lexical types s'amuser and amuser if in fact only one form, amuser, had been produced.

A few additional distinctions, not found in **le français fondamental,** were made in the present study: for example, separating causative faire (e.g. elle a fait bouillir la soupe) from other non-causative uses of faire, and distinguishing uses of a polysemous verb such as laisser as instances of one of two types, meaning roughly 'to allow' and 'to leave'.

As the list of verb items was compiled, each item was simultaneously coded as lexically acceptable or not. An item was considered lexically unacceptable (an error) if it contained a major malformation of the verb stem (e.g. étuliser instead of utiliser), was not a French verb (e.g. blower), or violated native-like lexical use in the context (e.g. 16108: le chat est très peur). Criteria were lenient, allowing a doubtful case to be scored as lexically acceptable (for full details see Appendix B, p. 66). Grammatical errors were disregarded; these included errors in associated pronouns (including reflexive pronouns), homophonous spelling errors, errors in tense and finiteness, and errors of agreement in number and gender.

Three researchers were involved in compiling and coding the list of verb items: two native speakers of French and one competent non-native French speaker. Most of the listing and coding was done by one of the native speakers. Following completion of the list, a reliability check was carried out by a second coder. Based on compositions from 15 immersion students and five native speakers, 97% agreement was found on the listing of items. Coding of lexical errors initially produced some major discrepancies between coders — of 31 errors noted in the same 28 compositions by one or other of two native-speaker coders, there was initial agreement on only eight errors. Following clarification of the criteria for determining lexical errors, these disagreements were resolved, mostly in favour of the original judgment of the principal coder. A final reliability check on the coding of lexical errors revealed a respectable correlation of .83 ($p<.01$) between coders with respect to numbers of errors produced by each of 13 students. Some discrepancy remained, however, in pinpointing specific errors: of 70 errors noted by one or both coders, there was agreement on 33 (47.1%).

3. MEASURES OF LEXICAL PROFICIENCY

Using the data base for immersion students described above, five different quantitative measures were developed with a view to determining appropriate indicators of the students' L2 lexical proficiency in the production of written text. These measures were then statistically compared in correlational analyses in order to examine the extent to which they were providing similar information about each student's lexical skills relative to the skills of other students.

The first measure can be labelled 'lexical error rate', while the other four are variations on the theme of lexical richness (Faerch, Haastrup and Phillipson 1984). They will be referred to respectively as 'number of lexical types', 'lexical variety', 'lexical specificity', and 'lexical sophistication' (Linnarud 1983).

3:1 Lexical Error Rate

A first step in developing this measure was to determine how it should be calculated to control for differences in the length of compositions written by the students. A student writing very long compositions, for example, might make the same number of errors as a student writing very short compositions, but the proportion of error would be quite different for the two individuals.

Three possibilities were considered as indicators of composition length: (a) the total number of verb items listed for each student, (b) the number of correct verb tokens listed for each student, and (c) the number of obligatory contexts for finite verbs produced by each student. A correlational analysis revealed that these three measures of length were very closely related, with intercorrelations of .98 between (a) and (b), .92 between (b) and (c), and .95 between (a) and (c). It was decided, therefore, to retain only measure (a) — number of verb items — since this was the measure that correlated best with each of the other two and that contained the most complete information.

Having determined an appropriate measure of composition length, a lexical error rate was arrived at for each student by dividing the number of lexical errors by the number of verb items in the compositions: for example, 5 errors ÷ 45 items = .11.

An alternative, 'equal-base' method of calculating the lexical error rate was also considered, in which length of composition was controlled by eliminating all verb items produced by each student above a certain cut-off point for each composition. These cut-off points were established by reference to the approximate minimum number of items produced by any student on a particular composition (see Table 1, p. 44), giving a total of 36 items for almost all students. The number of errors produced by each student up to the cut-off point was then divided by 36 (or number of items produced if less than 36) to produce the alternative measure of lexical error rate. A correlational analysis revealed that this measure was providing much the same information as the original measure of lexical error rate (Pearson correlation coefficient .87) in which the total number of verb items used by each student served as the base. It was decided to retain only the original measure based on all verb items since this contained more information than the equal-base method of calculation.

The measure of lexical error rate decided upon may thus be considered a relatively stable measure, affected little by the varying length of text produced by the students and even less by the various different ways of calculating length of text. A cautionary note needs, however, to be made with respect to content validity. Only those items that were considered lexically unacceptable by the lenient criteria of Appendix B were included in this measure, which thus probably represents a rather conservative estimate of the lexical errors that students were making. Another cautionary observation is that an individual student occasionally made the same type of error more than once. No effort was made to determine lexical error types, as opposed to error 'tokens', given the difficulty of determining what should count as the 'same' error. As has been pointed out in earlier studies of French immersion students' L2 production, (e.g. Harley and Swain 1978, 1984), outright errors are in any event a gradient phenomenon.

3:2 Number of Lexical Verb Types

This is the most straightforward of the measures of lexical richness that was calculated. It represents a count of each student's verb types that were identified in the master list, with 26 verb types removed that had occurred in the instructions to the compositions (see Table 2, p. 45). A problem with this measure, however, is that it does not take into account the length of text produced by each student. It can be predicted that students who produce longer texts will tend to use more different verb types, and this is in fact confirmed in the present study by a relatively high .68 correlation found between number of lexical verb types and length of text (i.e. the number of verb items listed for each student). While length of text may be partially dependent on prior vocabulary knowledge, it is clearly also dependent on a variety of other factors, such as linguistic skills of various kinds, speed of writing, and interest in the assigned topics. An indication of the effect of genre and topic on composition length, for example, is provided in Table 3, p. 46, which shows that on average the grade 6 students in the present sample tended to write longer narratives than request letters and within genres, to write more on animal topics than on other themes. While the topics were held constant for all students, and there was a significant tendency for those who wrote relatively long compositions on one topic also to be writing relatively long compositions on other topics (see Table 3), it is likely that some students were more stimulated by particular topics than were other students. In order to prevent such factors from affecting lexical proficiency scores, therefore, it was decided to develop other measures of lexical richness which would control for length of text.

3.3 Lexical Variety

This measure is the closest to the familiar 'type/token' ratio (see e.g. Faerch et al. 1984). In this instance, the verb types of each student (with the 26 verbs that were provided in instructions removed, as in 3:2 above) were controlled for length of text by dividing by the total of the verb items listed, including both lexically acceptable and lexically unacceptable items.

3:4 Lexical Specificity

This measure is similar to the above lexical variety measure, except that some further verb types are dropped: namely, those verbs that are reported by Muller (1974) as being among the 20 most frequent French verbs listed in each of three different sources: **le français fondamental** based on oral French and two frequency lists based on written French texts. These verbs (see Table 4, p. 47) are all of very general meaning, and they overlap in some instances with verbs provided in the instructions to the compositions. Their elimination from the list of types enables us to gauge the extent to which students are using verbs of less high coverage, or more specific meaning, in French. This reduced list of verb types is divided, as before, by the total of verb items to create the measure entitled 'lexical specificity'.

3:5 Lexical Sophistication

This is a measure of the extent to which a student is using relatively infrequent verbs. Only those verb types are selected from the data base that appear neither in **le français fondamental** (where roughly 200 verbs are listed) nor in the instructions to the compositions. This considerably reduced list of types is then divided by the total of verb items as in other measures.

3:6 Equal Base Measures of Richness

Figure 2, p. 61, summarizes the information contained in each of the above measures. As was the case for the measure of lexical errors, an alternative, equal base method of controlling for length of text, whereby a cut-off point of 36 items was used as a base for each student (see p. 26), was also examined with respect to the measures of lexical richness: number of verb types, lexical variety, lexical specificity, and lexical sophistication. Generally high correlations were found between the equal base method of calculation and the original measures: Pearson correlations were .60 for number of verb types, .74 for lexical variety, .78 for lexical specificity and .83 for lexical sophistication (all significant at the $p<.001$ level). As might be anticipated, the single original measure that did not embody a control for length of text — number of verb types — had the lowest correlation (.60) with the equal base method of calculation, where a control for length is built in. With respect to the other three measures of richness, it may be noted that when length of text is held constant, a strong association is maintained with the students' relative standings on the original measures. Only the original measures, therefore, which are based on the full data set and thus contain more information about each student, were maintained in further analyses.

3:7 Split-half Reliabilities

An effort was made to determine split-half reliabilities for the various measures of lexical proficiency described. These were calculated for the 68 immersion students by dividing the verb items listed for each student into 'odd item' and 'even item' subtests. Split-half reliabilities obtained for each of the lexical measures are provided in Table 5, p. 48. With the exception of number of lexical types, they are generally low, possibly because the verb item base for each subtest was too small and unstable. To test the hypothesis that size of data base was a factor in the low reliability coefficients, a second set of split-half reliabilities was calculated, including only those students who had produced at least 50 verb items. This made for two subtests based on a minimum of 25 items each. Despite the fact that the sample size was thereby reduced from 68 to 53, split-half reliabilities were improved for the three ratio measures of lexical richness — variety, specificity, and sophistication (see Table 5) — and especially for specificity and sophistication. This finding suggests that in working with such lexical measures, a substantial text base is necessary, and that the split-half measures obtained for the 68 students do not provide an adequate test of the reliability of the original lexical measures.

An alternative method of calculating split-half reliability was tried with the lexical sophistication variable, whereby four out of the five compositions were divided into two subtests, each containing a narrative and a letter (the bank robbery and the request to keep a puppy versus the rescue of a kitten and request to borrow a bicycle). The resulting α was .375 for the sample of 68 students, very close to the .362 obtained by the odd/even item method. When the sample was reduced to the same 53 students as before, a slightly lower α was obtained of .328. It should be noted that in this instance, a minimum of 50 verb items was not necessarily maintained for every student in the reduced sample, in that only four out of the five original compositions were included in the analysis. It was concluded that this method of calculating split-half reliabilities was also adversely affected by the problem of size of data-base.

3:8 Correlations Among Measures of Lexical Proficiency

Table 6, p. 48, shows the correlations obtained among the various measures of lexical proficiency described in 3:1-3:5 above. It indicates that there were positive relationships among all four measures of lexical richness, with lexical specificity showing the highest correlations with the other richness measures (.51 with number of verb types, .88 with lexical variety, and .61 with lexical sophistication, all significant at the $p < .01$ level). The lowest correlation among the richness measures was .25 ($p < .05$) between number of verb types and lexical sophistication.

In general, there was little relationship found between the measures of lexical richness and the lexical error rate. Only between the lexical variety measure and lexical error rate was there a significant correlation of -.34 ($p < .05$), indicating a slight tendency for those students with greater lexical variety to be making fewer lexical errors.

Two measures of lexical proficiency were selected for further use in a factor analysis (see Section 4 below). Lexical error rate was discarded owing to its skewed distribution and doubts about its meaningfulness as a measure of lexical competence (see 3:1 above). Also eliminated was the 'number of verb types' measure, since it did not incorporate any control for length of text. Of the remaining three measures of lexical richness, lexical specificity and lexical sophistication were retained in preference to lexical variety. The latter relatively easy measure of richness was seen as the least discriminating between 'good' and 'poor' vocabulary users in the present sample of L2 learners in that it incorporated highly frequent verbs such as *avoir*, *aller*, *dire*, *pouvoir*, whose use or non-use would tend to depend more on what a particular student had to say than on differential knowledge in relation to other students in the sample.

4. LEXICAL MEASURES AND L2 PROFICIENCY

In the theoretical framework designed to study the structure of L2 proficiency (Year 2 Report), three components of communicative competence were distinguished: grammatical competence, discourse competence, and sociolinguistic competence (see Figure 4, p. 63). It was hypothesized that these three types of competence ('traits') would be distinguished via factor analysis. Results of a confirmatory factor analysis presented in the Year 2 Report did not confirm the existence of the three traits as independent components of L2 communicative competence; instead a two-factor solution was found. The factors in this solution were identified as a general factor of second language proficiency, and a method factor that represented the context-reduced testing situation common to multiple choice tests and tests based on written compositions.

No special effort was made, at the time when the matrix of tests was developed, to design tasks that would focus specifically on aspects of lexical competence. However, it was suggested that vocabulary knowledge would enter, to some extent, into performance on all the tasks (Year 2 Report). An earlier position taken by Canale and Swain (1980), had been that vocabulary knowledge was part of grammatical competence.

4:1 Hypotheses

Arising from these two different perspectives on the relationship of vocabulary knowledge to L2 proficiency, the following hypotheses were of particular interest in the present study, using the grade 6 immersion data available from the original proficiency study in the Year 2 Report and incorporating the two new lexical measures — lexical specificity and sophistication — discussed in Section 3:

(1) That lexical competence is involved in all these components of language proficiency and that the lexical measures will therefore load on a general factor if there is one, or on all factors equally in an analysis without a general factor;

(2) That lexical competence is a part of grammatical competence, and that the lexical measures will therefore load most heavily on a grammatical factor (provided that there is one in the analysis).

Aside from these two mutually exclusive hypotheses, a third hypothesis, also incompatible with the other two, arises from preliminary analysis of second language data obtained from grade 3 French immersion students in another research project at the Ontario Institute for Studies in Education (Communicative French language testing: Consolidating and disseminating results of three recent evaluation projects in the MLC):

(3) That lexical competence is relatively distinct from the other components of language proficiency, and that the lexical measures will therefore constitute their own separate factor, or else have relatively high uniqueness by virtue of not sharing common variance with the other tests of language proficiency.

4:2 The Analysis

The starting point for the new analysis was a confirmatory factor analysis based on the nine proficiency measures of the earlier study, which led to the two-factor solution that was presented in the Year 2 Report. The lexical measures were added to this set of variables, and the analysis was repeated.[4]

The correlation matrix on which the analysis is based is shown in Table 7, p. 49. All correlations are Pearson product-moment correlations except those involving the discourse oral measure, which are polyserial correlations. Because the two lexical measures were somewhat similar, and based on the same data set, it was decided that the factor analysis should allow for correlated errors between these two variables. In fact, the three composition measures in the original set of nine measures were also based on partially overlapping data sets (see Figure 1, p. 60) although no provision was made for correlated errors in the original analysis. Therefore, a preliminary analysis was run that permitted correlated errors between all measures based on compositions. This analysis indicated that the correlations between errors for the two lexical measures, between errors for lexical sophistication and discourse composition, and between errors for discourse composition and sociolinguistic composition, were large enough to be important, but that the others were not. In order to minimize the number of free parameters, correlated errors were then allowed for these three pairs of tests only in the following analyses.

Two-factor analysis. The results of the two-factor analysis are shown in Table 8, p. 50. The chi-square of 43 degrees of freedom was 40.50 (p = .580), the adjusted goodness of fit index was .839, and the root mean square residual was .069. These measures of the overall fit of the model may be compared with those for the models to follow, and they show that the fit of the present model is acceptable. This seems to confirm hypothesis 1, since in this model lexical proficiency is considered to be part of a general language proficiency factor, represented here by factor 1. However, against hypothesis 1 is the fact that the communalities of the lexical measures in this solution are quite low — .11 for lexical specificity and .24 for lexical sophistication. Insofar as the lexical measures have variance in common with other variables in the set, that variance falls on the general language-proficiency factor; but most of their variance is unique. In fact, the loadings for lexical specificity are low and non-significant on both factors (.26 on the common factor, and .21 on the second factor which may be interpreted as a context-reduced written method factor), while the loading for lexical sophistication on the method factor is also low (.22). Lexical sophistication does have a significant loading on the common factor (.41).

Three-factor analysis. On the basis of results from a three-factor exploratory factor analysis which gave some indication of separate grammatical and discourse factors as well as a context-reduced written method factor, it was decided to try a confirmatory factor analysis with a three-factor model, in which the common factor of the above two-factor model was replaced by two factors, possibly correlated, representing grammatical and discourse competence respectively. This would allow us to test hypothesis 2 by looking at whether the lexical measures load more heavily on the grammatical or discourse factor. Since there is no sociolinguistic factor in this model, it is unclear where the sociolinguistic measures should load. It was decided on the basis of the loadings in the exploratory analysis to let the sociolinguistic oral and multiple choice measures load on the grammatical factor, and the sociolinguistic composition measure on the discourse factor.

In an initial run for the three-factor model, lexical measures were allowed to load on either the grammatical or discourse factors (as well as on the context-reduced method factor). However, the factor analysis did not converge for this model, and it was therefore re-run with the lexical measures allowed to load only on the grammatical factor (as well as on the method factor). Hypothesis 2 could still be examined because the factor analysis program provides 'modification indices', which show whether there would be a significant improvement if a given measure were allowed to load on another factor. In this case we examined whether there would be a significant improvement if the lexical measures were allowed to load on the discourse factor.

Results for this three-factor model are shown in Table 9, p. 51. The overall fit for this model is slightly better than for the two-factor model. Chi-square is 33.67 with 42 degrees of freedom (p = .817); adjusted goodness of fit index is .865; and the root mean square residual is .063. The correlation between the two trait factors is .53. The communalities of the lexical measures are similar to those for the two-factor model (.14 and .25 for specificity and sophistication, respectively), and the loadings on the grammatical factor for these measures are similar to or better than their loadings on the general factor before (.31 and .44, respectively; only the latter is significant). With regard to hypothesis 2, the modification indices show that there would not be a significant improvement in fit if the lexical measures were allowed to load on the discourse factor as well as on the grammatical factor. Hypothesis 2 is therefore supported, in that insofar as lexical measures load on trait factors, it is on the grammatical rather than the discourse factor. However, the large uniquenesses of the lexical measures remain in the three-factor model.

Four-factor analysis. In order to examine hypothesis 3, a four-factor model was tested, with the fourth factor being a separate lexical factor. Again this model was based on results from an exploratory analysis, where a solution with four factors was obtained in which the factors were identified as grammatical, discourse, context-reduced and lexical. Since the sociolinguistic oral measure loaded mainly on the lexical factor in the exploratory analysis, it was constrained to load on that factor in the confirmatory analysis as well. The outcome of this analysis is shown in Table 10, p. 52.

This model fits again somewhat better than the two- and three- factor models. The chi-square, with 39 degrees of freedom, is 29.21 (p = .873); the adjusted goodness of fit index is .875; and the root mean square residual is .059. The communality for lexical specificity remains quite small at .15, but the communality for sophistication increases to .43. This is reflected in the loadings of the lexical measures on the lexical factor: specificity is only .18, while sophistication is .52. This high loading for lexical sophistication on the lexical factor is not significant, however. The loadings of the lexical measures on the grammatical factor are reduced to .16 and .08 for specificity and sophistication, respectively.

4:3 Discussion

The three models that have been presented correspond to the three hypotheses that were to be tested, and all three fit reasonably well. Model 2 fits somewhat better than model 1, and model 3 fits somewhat better than model 2, but each successive model is also more complex than the previous one, and has more parameters, so that one must consider whether the gain in fit is worth the loss in simplicity. The overall fit of the three models by itself, therefore, provides no conclusive evidence favouring any one of the three hypotheses over the other two.

However, the low communalities of the lexical measures, especially in the two- and three-factor solutions, suggest that these measures do not share a great deal of variance with the other measures of language proficiency. These low communalities must be considered in light of the reliabilities of the lexical measures, which are not especially high, and it is possible that more reliable measures of lexical richness would correlate more highly with the other measures of proficiency obtained. Nevertheless, the low communalities do suggest that hypothesis 3 may be at least partially correct, and that lexical proficiency may be somewhat separate from other aspects of second language proficiency.

The second factor analysis provides some evidence as well in favour of hypothesis 2, because it shows that if language proficiency is factor-analysed into separate grammatical and discourse traits, then lexical measures, insofar as they fall on any of the factors, fall on the grammatical rather than the discourse factor. This supports hypothesis 2, that lexical competence is part of grammatical competence although with two conditions: first, not all of lexical competence seems to be captured by this grammatical component, since the communalities of the lexical measures remain low in the three-factor solution. And second, separate grammatical and discourse factors do not fit the present set of data very much better than a single general factor does, yet this two-trait model must be assumed in order to show that the lexical measures fall on the grammatical factor.

It is hypothesis 1 that receives the least confirmation from the present data. It is only in model 1 where there are no trait factors that lexical proficiency seems to be involved equally in all the aspects of proficiency examined. In the models that contain trait factors, models 2 and 3, one finds that the lexical measures go partly with grammar and partly as a separate factor and hence are not involved equally in the various aspects of language proficiency.

Apart from these results concerning the lexical measures, the factor analyses that have been presented are interesting because they show that grammatical and discourse factors can be found in the present data. Grammatical and discourse factors had not emerged in previous exploratory analyses involving only the nine non-lexical measures, although other evidence for their existence had been found (see Year 2 Report). These factors first emerged in an exploratory three-factor analysis, in which the new lexical measures were included, where the grammatical measures clearly clustered on one factor and the discourse measures on another. The third factor was interpreted as a context-reduced method factor.

This three-factor solution did not contain a separate sociolinguistic factor, and the loadings of these measures were therefore forced to fall on the existing grammatical and discourse factors. Sociolinguistic oral and sociolinguistic multiple choice loaded mostly on the grammatical factor which can be interpreted as indicating the importance of knowing the appropriate grammatical forms in these tests. Sociolinguistic composition loaded most heavily on the discourse factor, which can be interpreted as showing the importance of discourse skills in carrying out this composition task. These empirically determined loadings were then used to place the sociolinguistic tests on appropriate factors in the confirmatory three-factor analysis presented above.

In the same way, the four-factor confirmatory analysis was based on a corresponding exploratory analysis in which a lexical factor emerged. Again there was no sociolinguistic factor, and this time the sociolinguistic oral measure fell primarily on the lexical factor. This may be interpreted to indicate that there are various sociolinguistic 'formulae' that are important in this test that are learned as single chunks, interpretable as lexical units.

5. LEXICAL USE OF IMMERSION STUDENTS AND NATIVE SPEAKERS

In sections 3 and 4 above, measures of the lexical proficiency of immersion students have been examined in relation to each other and in relation to measures of other aspects of their language proficiency. In this section, the use of French verb vocabulary by the immersion students is compared with that of 22 native speakers of French at the same grade level. The purpose of this analysis is to gain insights into the nature of the lexical proficiency of immersion students.

5:1 Some Hypotheses

We would not expect the immersion students in this study to have as large or rich a vocabulary available for productive use in their compositions as the native speakers, given the relatively confined classroom context of their exposure to the second language. Based on hypotheses proposed by Levenston (1979) we can predict some ways in which these learners' verb vocabulary may be limited:

(1) they will prefer high-coverage verbs which can be generalized to use in a large number of contexts; and

(2) they will fail to use verbs that present problems of various kinds —semantic, syntactic, morphological, phonological or orthographic.

We may also predict that cross-linguistic influence (Kellerman 1984) from the L1 will be an important factor in the lexical use of the immersion students (e.g. Adjemian 1983, Blum-Kulka and Levenston 1983, Ijaz 1985, Ringbom 1978). Blum-Kulka and Levenston, for example, argue that all L2 learners make the initial assumption of word-for-word translation equivalence as a working hypothesis in dealing with the L2, and indeed that 'positive transfer' is an important way of increasing one's control of the L2 lexicon. Adjemian, in discussing the kinds of lexical rules and properties incorporated in recent versions of generative grammar, hypothesizes that L2 learners will make use of ready-made hypotheses based on L1 lexical rules or features wherever they perceive them to fit the available L2 data.[5] Adjemian (1983:254-5) is careful to point out that he is not "attempting to predict actual learner behavior but only potential learner behavior." With respect to immersion students, then, it may be predicted that:

(3) they will show a tendency not to use lexical verbs for which there is no direct translation equivalent in the L1; and

(4) they will project aspects of their L1 lexical knowledge onto their developing interlanguage lexicon.

These predictions are examined below.

5:2 Relative Frequencies of Verb Types

Table 11, p. 53, indicates that the native speakers produced on average significantly more text (in terms of number of verb items), more verb types, and more non-**français fondamental** verbs than did the immersion students. Table 11 also shows that the native speakers did significantly better on all the lexical measures in which length of text was controlled for: lexical error rate, variety, specificity, and sophistication. These initial comparisons thus indicate, as might be expected, that the lexical proficiency of the second language learners in the immersion classroom setting is not on a par with that of native speakers of the same age.

Notwithstanding these overall differences, Table 12, p. 54, shows that there is a significant tendency for the immersion students and the native speakers to use certain verbs with greater relative frequency than other verbs. The order of frequency with which the two groups use 158 high-coverage verbs from **le français fondamental** (excluding those provided in the instructions to the compositions) correlates at .64 ($p < .01$). When only those 22 very high frequency verbs of general meaning are considered which appear in the top twenty on several frequency lists (Muller 1974), an even higher correlation between the immersion students and native speakers is found (.881, $p < .0001$). These findings indicate that the immersion students, in dealing with the assigned composition topics, are quite similar to the native speakers in finding certain verbs of general meaning relatively more available or useful than others in expressing their ideas.

The relative frequencies with which both the native speakers and the immersion students use verbs from **le français fondamental** also correlate positively, but not as strongly, with the order found in the frequency list itself. The correlations with **le français fondamental** are very similar for the two groups: .42 ($p < .01$) for the immersion students and .43 ($p < .01$) for the native speakers, when 158 verbs are taken into account, .59 ($p < .001$) for the immersion students and .47 ($p < .01$) for the native speakers when only the 22 highest frequency verbs are taken into account.

The fact that the correlations with **le français fondamental** are not as high as they are between the immersion students and the native speakers is an indication of the importance of context, and also age, in considering the use of vocabulary by second language learners. **Le français fondamental** is a frequency list based on the oral vocabulary of adults across a wide range of conversational topics. In the present study, we are dealing with written vocabulary used by 11-12 year olds on five specific, assigned compositions. Clearly, there is nonetheless some relationship between **le français fondamental** frequencies and those found for the grade 6 immersion students and native speakers, suggesting that at least some French verbs are more readily used than others in a great variety of discourse contexts — by native speakers as well as L2 learners. It is important to note that the immersion students have not been exposed to French learning materials which are based on vocabulary specifically drawn from **le français fondamental**, as is the case for some French language texts designed for use in regular 'core' French classes. It would be of relevance, however, to determine the extent to which the immersion lexical frequencies correlate with those demonstrated in immersion teacher talk. Such an analysis could be done in the future from the grade 6 classroom transcripts used as the data base in Chapter 5, although these classes are not the same as those in the present study. Also of relevance is to consider the extent to which the high frequency French verbs listed in **le français fondamental** are matched by near-congruent high frequency verbs in English. West's (1953) **A General Service List of English Words** suggests that there are strong cross-linguistic similarities. Among the most frequent English verbs and modal auxiliaries listed by West, for example, are: have, say, can/could, make, go, come, see, take, know, like.

5:3 Lexical Repertoire of Immersion Students

A comparison of the number of different verb types produced by the 22 native speakers and a randomly selected subsample of 22 immersion students reveals that the immersion students are operating in their compositions with a much smaller stock of verb types. Table 13, p. 55, shows that the 22 immersion students used a combined total of 176 different verbs in their compositions, while the native speakers used 279, or 59% more verb types than the immersion students. While both groups included most of the verbs that were provided in the instructions to the compositions as well as the very high frequency verbs listed by Muller (1974), a major difference in the use of verbs that are not included in **le français fondamental** is evident. The 22 immersion students used 53 such 'sophisticated' verbs between them while the 22 native speakers, with a combined pool of 142, used more than 2 1/2 times as many different sophisticated verbs as the immersion students. The native speakers' pool of such verbs was not only larger in absolute terms but also larger in proportion to the total number of verb types in their combined production: 50.9% of the native speakers' verb types versus 30.1% of the immersion students' verb types. In keeping with our initial hypothesis, these findings show the second language learners as relatively more likely to select high frequency verb types than the native speakers while drawing on a smaller pool of verb types.

A further indication of the relatively restricted pool of verb types drawn upon by the 22 immersion students is the fact that only 38 verb types were used exclusively by these immersion students and not by the native speakers (see Table 13). All other verb types in the immersion pool of 176 were also used by one or more native speakers. In contrast, the number of verb types used exclusively by the native speakers was 141, including 108 verb types not listed in **le français fondamental** nor provided in the composition instructions (versus 19 such verbs used exclusively by the 22 immersion students).

5:4 Comparison of Immersion and Native-speaker Use of Verb Lexis

As already indicated (see 5:2 above), the immersion students were rated as having considerably more lexical errors in their compositions than were the native speakers, and they tended to make greater relative use of high frequency verbs. An analysis of the errors made by the immersion students provides some important clues as to the kinds of lexical problems they may be having, but it is equally important both in relation to L2 acquisition theory and from a practical diagnostic perspective to consider the ways in which their use of French verb lexis differs from that of native speakers without these differences necessarily resulting in errors.

In line with hypotheses 3 and 4, a distinct tendency can be observed for the immersion students to prefer those verb types in French that will fit into syntactic frames and argument-predicate structures that are congruent with English, and to avoid those verb types that have no direct, or relatively uncommon translation equivalents in English. Outright errors may be noted when verbs are selected by learners on the assumption that they are fully congruent with English with respect to the frames into which they will fit when in fact they are only partially congruent. In some cases, this can result in the wrong choice of verb (e.g. student no. 56108: le chat qui est très peur), in other cases, the error may be located elsewhere than in the verb itself (e.g. 16108: je vais au maison des Dupont et je demande pour une échelle). Note that pour in the latter example will have been scored as a grammatical, and not a lexical error, in the context of the present study.

Harley (in press) presents one example of a semantic domain where the immersion students demonstrate a preference for expressing the relevant information in a way that is congruent with English. In a study based on data from one of the narrative compositions in the present study, she shows that a subgroup of the immersion students are less likely to select verbs combining the notion of motion and direction than are native speakers, preferring often instead to express the notion of direction in a preposition phrase, as is characteristic of English (see e.g. Talmy 1975, Vinay & Darbelnet 1977). Further evidence of this pattern of use by immersion students is now available for the full sample of students in the present study. Table 14, p. 56, shows that the immersion students were making substantially less use than the native speakers of a number of common French verbs combining the notions of motion and direction: arriver, descendre, monter, partir, rentrer, and sortir. On the other hand, they were making as much, or considerably more use of several other verbs of motion which have direct translation equivalents and which in general can be fitted more readily into semantic and syntactic frames that are common in English. Among these verbs were the two highest frequency motion verbs in French, aller and venir. While their broad general meaning and high frequency in French may have been one factor in their selection by the immersion students, it is noteworthy that, in combination with preposition phrases or adverbials, they are sometimes pressed into service by the immersion students to express the more specific directional notions of e.g. sortir, rentrer, etc. along the lines of English come out, go in/home: e.g. 66114: Après un heure, Madame Dupont a venu au balcon, elle a vu le petit chat et criais "Oh mon petit Puff". Elle est allé dans la maison et téléphoné les pompiers). Of the remaining four motion verbs used as much, or more often, by the immersion students, three (courir, grimper, and sauter) have direct translation equivalents in English 'run', 'climb' and 'jump', while their relatively frequent use of the fourth, entrer, more explicitly directional in character, is in contradiction to their general tendency to make less use of such verbs than the native speakers. Perhaps in this case, positive transfer of the cognate 'enter' in English and the classroom use by teachers of "Entrez!" can be suggested as possible influences.

The apparent general tendency to seek translation equivalents in the second language that will fit the same kind of structural frame may be promoted by the existence of numerous cognates in the two languages. These can be a distinct asset to the learner. If we look, for example, at the 19 sophisticated verbs that were unique to the pool of verb types used by the subsample of 22 immersion students, we find that 14 of them have cognates or near-cognates in English (see Table 15, p. 57). Problems can also arise, of course, with false cognates or partially congruent items. An example in the narrative compositions is when approximately one third of the immersion used the verb chasser in the sense of 'to chase' in an inappropriate context, which was scored as an error by the native-speaking raters (e.g. 46120: les policiers chasse (sic) les bandits et met dans prison). On occasion, students are even led to invent cognates (e.g. 26102: il ne barque pas 'he doesn't bark'), on the assumption that there are not only regular semantic and syntactic correspondences between English and French lexical verbs but also phonological and orthographical ones.

Along with the multifaceted phenomenon of lexical transfer, the issue of inherent complexity within the second language needs also to be examined. Together these factors may combine to cause problems (hypothesis 2). Inherent complexity may, for example, have played a role in the differential use by immersion students and native speakers of some modal types of verbs in French, i.e. verbs that express semantic distinctions similar to those of the English modal auxiliaries (*can*, *may*, *must*, etc.). As can be seen from Table 16, p. 58, the native speakers tended in their request letters to make substantially more use than the immersion students of *aimer* and *vouloir* (e.g. 16204: *j'aimerais* vous demandez la permission d'utilisé la bicyclette qui est dan le garage; 16213: je *voudrais* que vous me la prèteriez jusque à la fin d'août). In these formulations of requests, it may be noted that the native speakers frequently used conditional forms of the verbs in question. The immersion students may have avoided using this type of request frame precisely because it would have involved the use of the relatively complex conditional verb form. Instead, the immersion students in making their requests were much more likely to use *laisser* in the sense of 'to allow' and *pouvoir*, often in combination with *s'il vous plaît*. In the letters asking to keep a puppy in an apartment and to borrow a bicycle, for example, the immersion students typically requested permission in the following manner: 36108: S'il vous plaît, *laisser* moi garder mon chien; 26127: Est-ce que je *peux* juste gardé pour 5 jours...; 66121: Alors, monsieur, est-ce que tu *peux* me prêter ta bicyclette? While the form of the immersion students' requests may in general have been inherently simpler from a grammatical point of view than those of the native speakers, this does not rule out the possibility that the greater use of *pouvoir* in a direct question frame also reflects a tendency in English to phrase such requests with *can/could*. Further data in the form of English compositions from the same-aged students are needed to examine this issue.

Inherent complexity may have combined with other factors to make the use of pronominal verbs rare among immersion students, even when credit is given for omitted reflexive pronouns (scored as grammatical errors). Thus, in addition to inherent morpho-syntactic complexity with respect to the use of pronouns and the marked auxiliary *être*, relative infrequency of use of pronominal verbs in the French of immersion teachers, together with lack of congruence with English structural frames, may have affected the pattern of immersion students' lexical use in this area. A count of pronominal verb types indicates that the native speakers used 52 such verbs between them, or 18.6% of their combined verb types, compared with only 16, or 9% of the verb types used by the subsample of 22 immersion students.

Apart from the suggested influence of morpho-syntactic complexity in the low use of 'polite' conditional forms and of pronominal verbs, inflectional complexity does not appear to have been a deterrent to the use of specific verb types. The subsample of 22 immersion students together used proportionately a slightly greater variety of 'irregular' verb types from the second and third groups (e.g. *finir*, *prendre*, *devoir*) than did the native speakers: 52:176, or 29.5% of total verb types, for the immersion students versus 65:279, or 23.3% of verb types in the native speaker group. The inflectional complexity of such verbs appears to have been reflected in conjugation errors (considered grammatical errors in the context of this study) rather than in the avoidance of specific verb types. It may be noted that many high coverage verbs in French are irregular in this sense. In fulfilling the expectations of hypothesis 1, therefore, the learners' lexical use is bound to be in some conflict with hypothesis 2.

Derivational complexity, on the other hand, appears to be a more useful predictor of lower use of certain verb types by the immersion students and the native speakers. However, since derivationally complex verbs tend also to be relatively low in frequency (note that there are no derived verbs in Muller's (1974) list), it is not possible to say whether frequency/utility or derivational complexity is the more important consideration in the relatively low use of such verbs. The subsample of 22 immersion students produced an estimated 25 derived verb types between them, representing 14.2% of their pool of 176 verb types, while the native speakers produced 54 derived verb types, or 19.3% of their pool of 279 verb types.[6] If derived verbs are divided into (a) those that are derived from other verb forms by affixation (e.g. *apporter*, *disparaître*, *emporter*, *ramener*), and (b) those that are derived from other parts of speech such as nouns or adjectives with or without affixation (e.g. *affoler*, *encercler*, *paniquer*), then it appears that the immersion students are more likely to use (a) than (b). Nineteen out of the 25 derivationally complex verb types that they used (76%) may be considered derived by affixation from other verb forms; among the native speakers, the proportion was 33:54, or 61%. Of the six verbs used by immersion students which can be considered derived from other parts of speech, five have cognates in English (*approcher*/'approach', *garder*/'guard', *masquer*/'mask', *respecter*/'respect', *téléphoner*/'telephone',). These findings suggest that the immersion students are transferring such items from their L1 and making little use of L2-internal derivational resources of the kind that link different parts of speech. In order to determine whether the immersion students' patterns of affixation were similar to those of the native speakers, verb prefixes used with more than one verb type in the native speaker group's compositions were examined. Table 17 (p. 59) shows that *re-/r-*is the most widely distributed form of prefix, followed by *en-*, then *a-*, *é-* and *dé-*. Among the subsample of 22 immersion students, *re-/r-* is also the most widely distributed prefix, followed by *a-*, *dé-* and *é-*. There is no use by these immersion students of the prefix *en-*. One way in which the lexical resources of immersion students could be enlarged is by focussing attention in the classroom on derivational processes of affixation, perhaps with particular emphasis on those which appear currently to be underused.

In sum, this brief comparative analysis of verb lexis used by the immersion students and the native speakers provides general support for the hypotheses presented in section 5:1. The immersion students tended to make proportionately more use of high-coverage verbs than did the native French speakers (hypothesis 1), and to make less use of some syntactically and morphologically complex verbs such as pronominal and derived verb types (hypothesis 2). It may be noted that, in line with hypothesis 1, these complex verbs tend not to be among the most frequent, high-coverage verbs in French. Inflectional complexity, on the other hand, which is characteristic of a number of high frequency verbs in French, does not in general appear to have been a deterrent to their use by the immersion students (with the possible exception of request verbs that might have entailed the use of conditional forms). Semantic and/or syntactic incongruence with English L1 is clearly an important factor in the immersion students' non-use of some L2 verb types and in many of their lexical (as well as grammatical) errors (hypotheses 3 and 4). At the same time, there is evidence to suggest that lexical similarities between the L1 and the L2, specifically in the form of cognates, help to increase the students' L2 lexical resources.

6. CONCLUSION

The present study has compared several different measures of immersion students' lexical proficiency as demonstrated in their productive use of verb lexemes in written compositions. Two measures of lexical richness — specificity and sophistication — were selected for use in a factor-analytic study. These were chosen in preference to a measure of lexical errors, given the lack of agreement by native-speaker judges in rating of errors. Clearly other kinds of lexical measures, based on different L2 tasks, are also needed to gain a fuller understanding of the students' lexical competence. In particular, little is yet known about the receptive lexical proficiency of immersion sudents.

Based on the factor analyses described in section 4, no strong conclusions can be drawn concerning the relationship between lexical richness and other aspects of immersion students' proficiency in French, although there is at least some evidence to suggest that there are aspects of lexical proficiency that represent a somewhat separate dimension of L2 proficiency for these students. An interesting additional finding was that grammatical and discourse factors emerged in the three factor solution that had not previously been found in the main proficiency study (Year 2 Report).

The descriptive comparison of lexical use by the immersion students and native speakers of French revealed different patterns of use that were in line with general predictions that had previously been made concerning L2 lexical development. The adherence of immersion students to mostly high coverage verbs and patterns of lexical use that are congruent with their English L1 suggest a number of implications for classroom practice. The students' stock of available lexical items might, for example, benefit from more activities designed to increase their use of derivational processes in the second language, and from an emphasis on the use of specific terms rather than more general, high-coverage terms. An explicit classroom focus on particular semantic domains, such as directional expressions, where the L1-influenced lexical use of the immersion students tends to differ systematically from that of the native speakers, might also be beneficial. In short, there appears to be a need for further examination of different aspects of the lexical proficiency of immersion students, and for classroom experimentation involving a variety of approaches to the teaching of lexical competence.

Footnotes

[1] The program in which these students are enrolled consists of 100% schooling in French from kindergarten to grade 2, following which a period of English language arts is introduced in grade 3. The English portion of the day is gradually increased so that by grade 6 the students are receiving about 50% of their schooling in French and the other 50% in English.

[2] Given the difficulty in determining, for specific cases, whether such constructions should be interpreted as être + adjective, or as resultative auxiliary (être) + participle, it was decided to include all such constructions as if they were auxiliary + past participle.

[3] Note that in these analyses, the original sample of 69 immersion students is reduced by 1 to 68, owing to the elimination of a student who was found to be an outlier on several critical variables.

[4] At the same time an accent measure was added, and is incorporated in Tables 7 - 10. It took the form of a 4-point rating scale based on data from the original grammatical oral measure. Since this variable was not directly relevant to the present study of lexical competence, it is not considered further.

[5] Based on Kellerman (1979), Adjemian is allowing for the possibility that learners' general perceptions of language relatedness may influence their willingness to transfer lexical items.

[6] Verbs were listed as derived if they contained a non-inflectional affix (e.g. disparaître) or were derived from another part of speech (e.g. téléphoner). The designation of 'derived' was based on a consensus of two researchers with linguistic training.

References

Adjemian, C. The transferability of lexical properties. In S. Gass and L. Selinker (eds.) **Language Transfer in Language Learning**. Rowley, Mass.: Newbury House, 1983.

Allen, J.P.B., Cummins, J., Mougeon, R. and Swain, M. **The Development of Bilingual Proficiency: Second Year Report**. Toronto: OISE, 1983 (mimeo).

Blum-Kulka, S. and Levenston, E.A. Universals of lexical simplification. In C. Faerch and G. Kasper (eds.) **Strategies in Interlanguage Communication**. London: Longman, 1983.

Canale, M. On some dimensions of language proficiency. In J. W. Oller (ed.) **Issues in Language Testing Research**. Rowley, Mass.: Newbury House, 1983.

Canale, M. and Swain, M. Theoretical bases of communicative approaches to second language teaching and testing. **Applied Linguistics**, 1980, 1(1), 1-47.

Faerch, C., Haastrup, K. and Phillipson, R. **Learner Language and Language Learning**. Clevedon, Avon: Multilingual Matters, 1984.

Faerch, C. and Kasper, G. (eds.) **Strategies in Interlanguage Communication**. London: Longman, 1983.

Gougenheim, G., Michéa, R., Rivenc, P. and Sauvageot, A. **L'Elaboration du français fondamental, 1er degré**. Philadelphia, Chilton, 1964.

Harley, B. Transfer in the written compositions of French immersion students. In H. W, Dechert and M. Raupach (eds.) **Transfer in Production**. New York: Ablex, in press.

Harley, B. and Swain, M. An analysis of the verb system used by young learners of French. **Interlanguage Studies Bulletin**, 1978, 3(1), 35-79.

Ijaz, H. Native language and cognitive constraints on the meaning ascribed to select English spatial prepositions by advanced adult second language learners of English. Ph.D. thesis, University of Toronto, 1985.

Johansson, S. Studies of error gravity. Native reactions to errors produced by Swedish learners of English. **Gothenburg Studies in English**, 1978, 44.

Juillard, A., Brodin, D and Davidovitch, C. **Frequency Dictionary of French Words**. La Haye, 1970.

Kellerman, E. Transfer and non-transfer: Where we are now. **Studies in Second Language Acquisition**, 1979, 2(1), 37-57.

Kellerman, E. The empirical evidence for the influence of L1 in interlanguage. In A Davies, C. Criper, and A.P.R. Howatt (eds.) **Interlanguage**. Edinburgh: University of Edinburgh Press, 1984.

Lepicq, D. Aspects théoriques et empiriques de l'acceptabilité linguistique: Le cas du français des élèves des classes d'immersion. Ph.D. thesis, University of Toronto, 1980.

Levenston, E.A. Second language lexical acquisition: Issues and problems. **Interlanguage Studies Bulletin**, 1979, **4**, 147-160.

Linnarud, M. On lexis: The Swedish learner and the native speakers compared. In K. Sajavaara (ed.) **Cross-language Analysis and Second Language Acquisition. Vol. 2.** Jyväskylä: University of Jyväskylä, 1983.

Muller, C. Les verbes les plus fréquents du français. **Le Français dans le monde**, 1974, **103**, 14-17.

Politzer, R.L. Errors of English speakers of German as perceived and evaluated by German natives. **Modern Language Journal**, 1978, **62**, 253-261.

Ringbom, H. The influence of the mother tongue on the translation of lexical items. **Interlanguage Studies Bulletin**, 1978, **3**(1), 80-101.

Shapira, R.G. The non-learning of English: Case study of an adult. In E. M. Hatch (ed.) **Second Language Acquisition: A Book of Readings.** Rowley, Mass.: Newbury House, 1978.

Swain, M. Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden (eds.) **Input in Second Language Acquisition.** Rowley, Mass.: Newbury House, 1985.

Talmy, L. Semantics and syntax of motion. In J. Kimball (ed.) **Syntax and Semantics Vol. 4.** New York: Academic Press, 1975.

Vinay, J.P. and Darbelnet, J. **Stylistique comparée du français et de l'anglais.** Montreal: Beauchemin, 1977.

West, M. **A General Service List of English Words.** London: Longman's Green, 1953.

Table 1

Cut-off Points for Equal-base Method of Assessing Lexical Proficiency

Composition	Cut-off point
Narr. 1 - rescue of kitten	10
Narr. 2 - bank robbery	8
Letter 1 - keeping puppy	8
Letter 2 - borrowing bicycle	7
Letter 3 - pollution project	3

Table 2

Verb Types Provided in Instructions to Compositions and Eliminated from Measures of Lexical Richness

- barrer
- commencer
- convaincre
- (se) demander
- déménager
- donner
- dormir
- écrire
- envoyer
- être
- (se) faire
- garder
- (s') imaginer
- louer (location)
- (se) présenter
- (se) servir
- suivre
- travailler
- (se) trouver
- utiliser
- voir
- vouloir

Table 3

Comparison of Length of Five Compositions
Written by 68 Immersion Students

A. Analysis of variance

Source		SS	MS	F	p
Between compositions	4	2750.21	687.55	79.72	.001
Error	268	2311.39	8.62		

B. Multiple comparisons

Comparison		Value of comparison
Narratives:	rescue of kitten vs. bank robbery	2.22*
Letters:	keeping puppy vs. borrowing bicycle	1.28
Topics:	Animals vs. other	3.50*
Genre:	letters vs. narratives	4.21*

* Comparison greater than the critical for Scheffé, $\alpha = .05$

C. Correlations in length

	Bank Robbery	Keeping Puppy	Rescue of kitten	Borrowing bicycle
Keeping puppy	.31**			
Rescue of kitten	.34**	.37**		
Borrowing bicycle	.33**	.41**	.36**	
Info. on pollution	.23*	.20*	.24*	.36**

* $p < .05$ ** $p < .01$

Table 4*

High frequency verbs in French (in declining order of frequency) deleted from numerator of lexical specificity measure

	a	b	c
être	1	1	1
avoir	2	2	2
faire	3	3	3
dire	4	5	4
pouvoir	5	4	8
aller	6	6	5
voir	7	7	6
savoir	8	8	7
vouloir	9	9	10
venir	10	10	11
falloir	11	12	9
devoir	12	11	17
croire	13	15	14
trouver	14	14	19
donner	15	13	20
prendre	16	16	12
parler	17	17	18
aimer	18	20	26
passer	19	19	16
mettre	20	21	15
tenir	22	18	25
arriver	33	29	13

*Adapted from Muller (1974).

a Trésor de la langue française

b Frequency Dictionary of French Words (Juillard, Brodin & Davidovitch 1970) — written corpus

c le français fondamental (Gougenheim, Michéa, Rivenc & Sauvageot 1964) -- oral corpus

Table 5

Split-half Reliabilities (α) Based on Odd-item/Even-item Scores on Measures of Lexical Proficiency

Measure	Reliability	
	N = 68	N = 53[a]
Lexical error rate	.308	.278
No. of verb types	.798	.723
Lexical variety	.375	.420
Lexical specificity	.274	.516
Lexical sophistication	.362	.572

[a] Those students who produced at least 50 verb items across the five composition tasks.

Table 6

Correlations Among Lexical Proficiency Measures

	Lexical error rate	Verb types	Lexical variety	Lexical specificity
Verb types	-0.18			
Lexical variety	-0.34*	.43**		
Lexical specificity	-0.27	.51**	.88**	
Lexical sophistication	-0.13	.25*	.45**	.61**

* $p < .05$ ** $p < .01$

Table 7

Correlations of All Measures

	GO	GM	GC	DO	DM	DC	SO	SM	SC	SPEC	SOPH	ACCT.
GO	1.000											
GM	0.203	1.000										
GC	0.317	0.594	1.000									
DO	0.127	0.032	0.085	1.000								
DM	0.241	0.426	0.403	0.250	1.000							
DC	0.233	0.469	0.503	0.256	0.462	1.000						
SO	0.069	0.070	0.228	0.124	0.098	0.014	1.000					
SM	0.345	0.336	0.441	0.146	0.285	0.331	0.160	1.000				
SC	0.003	0.206	0.142	0.170	0.355	0.531	-0.043	0.098	1.000			
SPEC	0.046	0.292	0.315	0.120	0.057	0.125	0.127	0.128	0.072	1.000		
SOPH	0.126	0.393	0.322	0.196	0.165	0.015	0.327	0.177	-0.083	0.605	1.000	
ACCT	0.329	0.189	0.216	0.196	0.143	0.148	0.145	0.243	-0.015	0.273	0.263	1.000

Key to

GO	grammatical oral	SO	sociolinguistic oral
GM	grammatical multiple choice	SM	sociolinguistic multiple choice
GC	grammatical composition	SC	sociolinguistic composition
DO	discourse oral	SPEC	lexical specificity
DM	discourse multiple choice	SOPH	lexical sophistication
DC	discourse composition	ACCT	accent

Table 8

Confirmatory Factor Analysis, 2-Factor Solution

	Factor 1 General	Factor 2 Written	Communalities
GO	.55	—	.30
GM	.35	.71	.63
GC	.53	.57	.60
DO	.32	—	.10
DM	.34	.45	.32
DC	.37	.53	.44
SO	.33	—	.11
SM	.52	.24	.33
SC	.00	.33	.11
SPEC	.26	.21	.11
SOPH	.41	.22	.24
ACCT	.51	—	.26

Chi-square with 43 degrees of freedom = 40.50, p = .58
Adjusted goodness of fit index = .839
Root mean square residual = .069

Table 9

Confirmatory Factor Analysis, 3-Factor Solution

	Factor 1 Grammar	Factor 2 Discourse	Factor 3 Written	Communalities
GO	.52	--	--	.27
GM	.31	--	.73	.63
GC	.48	--	.63	.62
DO	--	.62	--	.38
DM	--	.46	.46	.41
DC	--	.48	.57	.57
SO	.35	--	--	.12
SM	.48	--	.28	.32
SC	--	.28	.28	.15
SPEC	.31	--	.19	.14
SOPH	.44	--	.22	.25
ACCT	.55	--	--	.30

Correlation between factors 1 and 2 = .53

Chi-square with 42 degrees of freedom = 33.67, p = .82
Adjusted goodness of fit index = .865
Root mean square residual = .063

Table 10

Confirmatory Factor Analysis, 4-Factor Solution

	Factor 1 Grammar	Factor 2 Discourse	Factor 3 Lexis	Factor 4 Written	Communalities
GO	.58	—	—	—	.34
GM	.29	—	—	.81	.75
GC	.52	—	—	.55	.59
DO	—	.46	—	—	.21
DM	—	.51	—	.39	.41
DC	—	.62	—	.46	.60
SO	—	—	.58	—	.33
SM	.55	—	—	.21	.36
SC	—	.37	—	.21	.18
SPEC	.16	—	.18	.25	.15
SOPH	.08	—	.52	.33	.43
ACCT	.49	—	—	—	.24

Correlation between factors 1 and 2 = .60
Correlation between factors 1 and 3 = .41

Chi-square with 39 degrees of freedom = 29.21, p = .87
Adjusted goodness of fit index = .875
Root mean square residual = .059

Table 11

Comparisons Between Immersion Student and Native Speaker Means on Lexical Measures

Variable	Group	$\overline{X}$	SD	T	p
No. of verb items	Imm	59.56	11.90	-3.43	.001
	Nat	70.64	16.63		
No. of verb types	Imm	25.41	5.46	-8.09	.000
	Nat	37.86	8.38		
Lexical error rate	Imm	0.06	0.04	5.73	.000
	Nat	0.01	0.02		
Lexical variety	Imm	0.43	0.06	-7.14	.000
	Nat	0.54	0.08		
Lexical specificity	Imm	0.29	0.06	-7.01	.000
	Nat	0.40	0.08		
Lexical sophistication	Imm	0.08	0.03	-7.38	.000
	Nat	0.17	0.06		

Table 12

Spearman Correlations of Use of Verbs by Immersion Students and Native Speakers with Frequencies in le français fondamental and Muller's List of High Frequency Verbs

	N	FF[a]	Imm[a]
Imm	69	.416**	
Nat	22	.428**	.640**

	N	Muller list[b]	Imm[b]
Imm	69	.59***	
Nat	22	.48**	.88**

[a] Based on those verbs used by either group and also listed in le français fondamental, excluding verbs provided in instructions to compositions.

[b] Not excluding verbs provided in instructions.

** $p < .01$

*** $p < .001$

Table 13

Verb types generated by 22 Immersion students and 22 Native speakers

	All types	Excluding provided verbs[a]	Excluding provided & high frequency verbs[b]	Excluding le français fondamental and provided verbs[c]
Total pool:				
Imm or Nat	317	294	271	161
Imm Total	176	154	134	53
Nat Total	279	257	235	142
Overlap	138	117	98	34
Imm Unique	38	37	36	19
Nat Unique	141	140	137	108

[a] Numerator for Variety

[b] Numerator for Specificity

[c] Numerator for Sophistication

Table 14

Mean Use per Student of Some Verbs of Motion in Two Narrative Compositions[a]

	Native speakers (N = 22)	Immersion students (N = 69)
Verbs used much more often by native speakers		
arriver	.96	.28
descendre	.82	.43
monter	1.09	.30
partir	.73	.30
rentrer	.78	.13
sortir	.77	.35
Verbs used as much or more often by immersion students		
aller	.87	.88
courir	1.23	1.78
entrer	.32	.65
grimper	.55	.55
sauter	.32	.68
venir	.59	.95

[a] Includes only verbs for which $\overline{X}$ use is at least .33 for one of the groups. Those items rated as errors are not included.

Table 15

Cognates and Partial Cognates Among the 19 'Unique' Sophisticated Verbs of 22 Immersion Students

French	English
adorer	adore
alarmer	alarm
apparaître	appear
apprécier	appreciate
avoir la permission	(have) permission
commander	-
découper	-
dégonfler	-
demeurer	-
démolir	demolish
garder[a]	guard
isoler	isolate
masquer	mask
punir	punish
se rapprocher	approach
rattacher	attach
respecter	respect
situer	situate
tressaillir	-

[a] In the sense of surveiller

Table 16

Mean Use per Student of Some Modal Types of Verbs in Request Letters[a]

	Native speakers (N = 22)			Immersion students (N = 69)		
	Keeping puppy	Borrowing bicycle	Info. on pollution	Keeping puppy	Borrowing bicycle	Info. on pollution
aimer	.36	.32	.64	.03	.09	.04
laisser (permission)	.27	.05	.00	.67	.30	.00
pouvoir (permission)	.36	.32	.00	.61	.77	.00
pouvoir (total)	.64	.86	.46	1.03	1.14	.97
vouloir	.77	.68	.36	.46	.29	.14

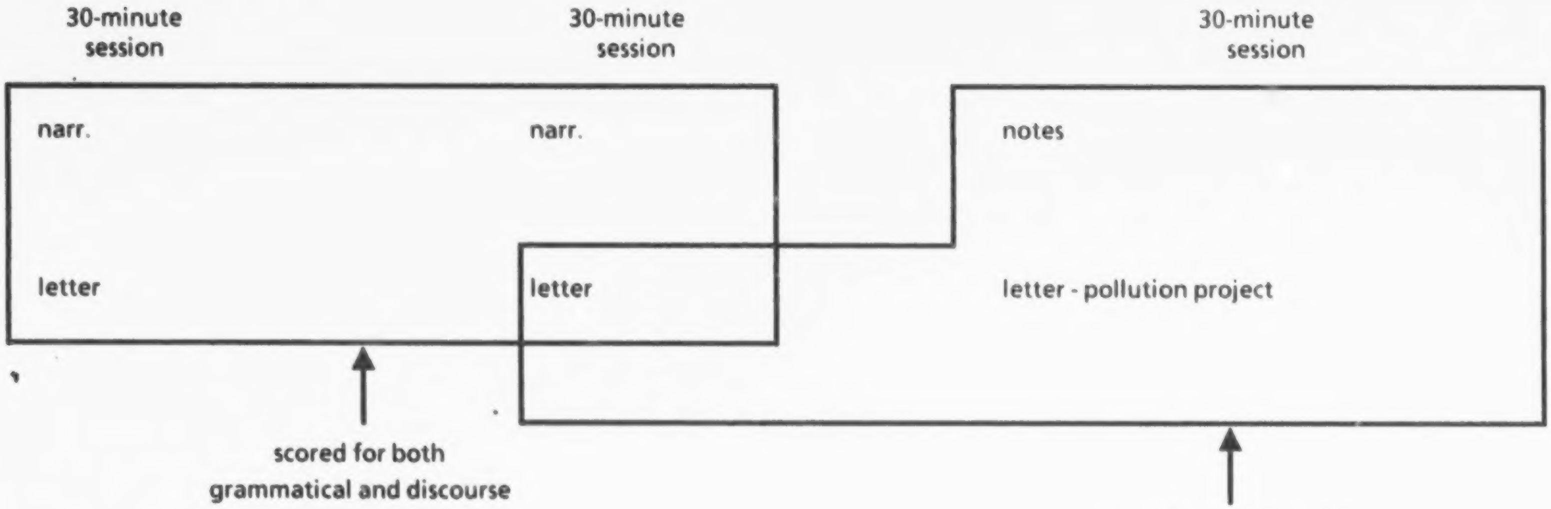
[a] Not including those items rated as errors

Table 17

Distribution of Derivational Prefixes Across Different Verb Types Among Immersion Students and Native Speakers

Prefix	No. of different verb types with given prefix	
	Imm (N = 22)	N Spkrs (N=22)
re-/r-	9	15
en-/em-[a]	0	9
a-	4	6
dé-	3	3
é-	1	3
Total	15	36

[a] The form en-/em- did not always have the same meaning


30-minute session
30-minute session
30-minute session
narr.
narr.
notes
letter
letter
letter - pollution project

scored for both grammatical and discourse competence

(order of administration of composition topics counterbalanced within and across testing sessions)

scored for sociolinguistic competence

(order of administration of notes and pollution project letter counterbalanced)

Figure 1
Scoring of Written Composition Tasks in Original Study (Allen et al. 1983)

1. **Lexical error rate**

$$\frac{\text{lex. errors in verbs}}{\text{all verb items}}$$

2. **Lexical types**

$$\text{verb types} - \text{verbs in instructions}$$

3. **Lexical variety**

$$\frac{\text{verb types} - \text{verbs in instructions}}{\text{all verb items}}$$

4. **Lexical specificity**

$$\frac{\text{verb types} - \begin{array}{l}\text{verbs in instructions}\\ \text{hi-frequency verbs}\end{array}}{\text{all verb items}}$$

5. **Lexical sophistication**

$$\frac{\text{verb types} - \begin{array}{l}\text{verbs in instructions}\\ \text{verbs in }\textbf{le français fondamental}\end{array}}{\text{all verb items}}$$

Figure 2

Calculation of Lexical Proficiency Measures

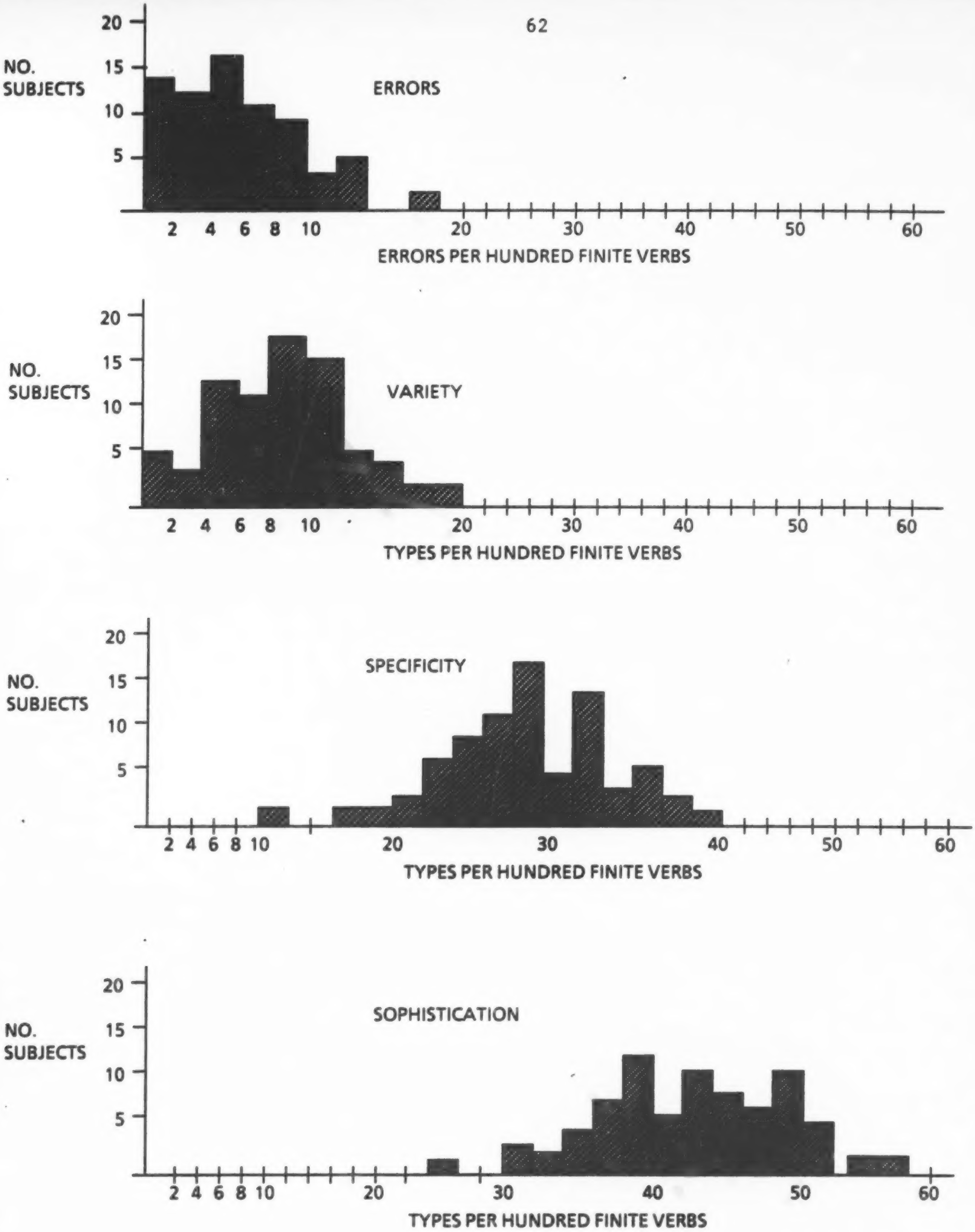


Figure 3

Distribution of Scores on Lexical Measures

	Traits		
Methods	Grammatical	Discourse	Socio-linguistic
Oral production	1,1 Interview	1,2 Movie retelling and discussion	1,3 Slide-tape situations
Multiple choice	2,1 Test de Grammaire	2,2 Discours à Choix Multiple	2,3 Socio-linguistique
Written production	3,1 Rédaction française Narrative: Au secours Letter: La bicyclette	3,2 Rédaction française Narrative: Aux voleurs Letter: Le chien	3,3 Ecrire en français

Figure 4

Matrix of Traits and Methods and Tests Used to Assess Each

Appendix A

Instructions for Compositions

Narratives:

Au secours!

Écris une petite histoire sur le sauvetage d'un petit chat se trouvant dans le haut d'un arbre.

Commence avec le début suivant:

C'était un beau dimanche d'été et sur le balcon de la maison des Dupont un petit chat dormait tranquillement. Tout d'un coup trois chiens ______________________________

Aux voleurs!

Écris une petite histoire sur un important vol de banque. Commence avec le début suivant:

Ce jour-là, comme d'habitude, la banque était pleine de monde. Tout d'un coup, trois bandits ____

Letters:

Imagine-toi que ta famille déménage et que le nouveau propriétaire ne veut pas que tu gardes ton petit chien dans votre nouvel appartement.

Écris une lettre au propriétaire pour le convaincre de te donner la permission de garder ton petit chien.

Sers-toi de l'espace ci-dessous:

Cher monsieur, ____

Imagine-toi que ta famille loue une maison à la campagne pendant le mois d'août. Dans le garage, qui est barré, tu vois une belle bicyclette dix-vitesses.

Écris une lettre au propriétaire de la maison pour le convaincre de te donner la permission d'utiliser la bicyclette.

Sers-toi de l'espace ci-dessous:

Cher monsieur, ____

Imagine que tu travailles à un projet sur la pollution. Dans deux semaines, ça sera à toi de présenter ton projet à la classe. Tu as été voir à la bibliothèque mais tu n'as pas trouvé le matériel nécessaire. Écris une lettre au Service de l'environnement pour leur demander de t'envoyer en urgence des informations et des photos pour faire ta présentation.

Appendix B

Criteria for Judging Lexical Errors

A. The following are not counted as lexical errors; the verb is listed in its correct form:

1. homophonous errors in the stem (e.g., miowler for miauler: listed as miauler)

2. grammatical errors such as:

   errors in gender, number, or tense (e.g., elle sont sorti);

   omission, substitution, or addition of preposition or reflexive pronoun (e.g., ils vont for ils s'en vont: listed as s'en aller; il le vient faire for il vient de le faire);

   use of an incorrect but recognizable form of an irregular verb (e.g., je vais le mis for . . . mettre;

   use of the incorrect auxiliary (e.g., il a passé for il est passé).

B. The following are counted as lexical errors:

1. major morphological errors in the stem (e.g. miower for miauler; protecter for protéger); the scorer should be lenient, but not beyond his or her tolerance as a native speaker; the stem should remain essentially French in form to be considered correct;

2. lexical errors in the use of the verb in context (e.g., dire for parler; chasser for poursuivre; prêter for emprunter).

   If there is no evidence of error (for example, if it is possible in the context that the student meant chasser in the correct French sense), the verb is accepted as correct.

Chapter 2

THE DEVELOPMENTAL GROWTH OF METAPHOR COMPREHENSION IN CHILDREN'S FIRST AND SECOND LANGUAGE

Janice Johnson

The study examines the development of metaphor comprehension in minority (Spanish-English) and majority (English) language children in Canada.[1] Previous work with English-speaking school-aged children has resulted in development of a coding method and a measurement scale for assessing the degree of conceptual sophistication in children's spontaneous metaphor interpretations (Johnson 1982, Johnson and Pascual-Leone 1984). Using this method, level of metaphoric understanding has been found to be largely determined by the child's mental-attentional capacity (working memory, general level of cognitive development), and to a much lesser extent by the child's level of linguistic sophistication or general knowledge structures. The present study examines metaphor comprehension in developmental samples who differ widely in English proficiency, but are equated on mental capacity. It thus makes a stringent test of the claim that metaphoric processing is constrained more by mental-attentional capacity than by linguistic sophistication.

In the study "linguistic proficiency" is operationalized in terms of traditional measures: The main measure is a standardized test of oral language proficiency (Woodock 1980, 1981); secondary measures are a story-retelling task and teacher ratings of proficiency (DeAvila and Duncan 1983). Although the metaphor interpretation task is not referred to as a "proficiency measure", it clearly does reflect a kind of proficiency, that is, the degree to which the child can use the first or second language for conceptual (i.e., metaphoric) processing at the level of abstraction appropriate for his or her level of cognitive development. Grammatical correctness is not important in the metaphor task, and it is not necessary to use complex vocabulary to convey a conceptually sophisticated understanding. The development of one's conceptual repertoire is likely facilitated by language, but is not equated with language. In the study, "level of cognitive development" is operationalized in terms of a content-free mental-attentional capacity that grows with age; this capacity is measured with a nonverbal task which has been shown to have good validity across cultural groups. A major goal of the study is to tease apart constraints that linguistic proficiency and cognitive development separately place on children's level of metaphoric interpretation.

Prior to a description of the study, some relevant issues from the field of second language learning are discussed.

1. THE DEVELOPMENT OF BILINGUALISM

Cummins (e.g. 1984b) has proposed that "language proficiency" can be conceptualized along two orthogonal dimensions. One dimension, which he labels "context-embedded vs. context-reduced", relates to the amount of contextual support available for expressing or receiving meaning. In context-embedded communication the participants can actively negotiate meaning and the language is supported by situational cues. By contrast, context-reduced communication relies primarily on linguistic cues to meaning. An example might be carrying on a conversation (embedded) vs. writing a composition (reduced).

The second dimension concerns the extent to which the linguistic task is cognitively demanding vs. cognitively undemanding. A cognitively undemanding communicative task is one in which the linguistic tools have become largely automatized or overlearned — such a task requires relatively little active cognitive involvement or mental effort. Cognitively demanding communication tasks, on the other hand, involve linguistic tools that have not been overlearned and, therefore, require active cognitive involvement. Cummins (1984b) conceptualizes cognitive involvement "in terms of the amount of information that must be processed simultaneously or in close succession by the individual in order to carry out the activity." It is this second dimension that seems most related to cognitive developmental aspects of language proficiency, and it is this dimension which is of most concern here. Cummins (1984b) has proposed that "any language task which is cognitively-demanding for a group of individuals is likely to show a moderate degree of interdependence across languages" (p. 14).

A related issue concerns the theoretical assumption of "separate underlying proficiency" vs. "common underlying proficiency" in bilingualism (Cummins, Swain, Nakajima, Handscombe, Green, and Tran 1984). The first position assumes that proficiency in L1 and L2 are separate; the second assumes that L1 and L2 proficiency are interdependent. As Cummins et al. (1984) point out, these assumptions have important implications for bilingual education, because "if L1 and L2 proficiency are manifestations of a common underlying proficiency, then instruction in either language is, theoretically, capable of promoting the proficiency underlying academic proficiency in both languages" (p. 61). Cummins et al. (1984) present evidence supporting the interdependence position, particularly in the case of cognitively demanding linguistic tasks.

A different, although similar, dimensionalization of linguistic skills has been proposed by Bialystok and Ryan (1985). Bialystok and Ryan propose the dimension of "high vs. low analyzed knowledge" and the dimension of "high vs. low cognitive control". The analyzed knowledge dimension refers to the extent to which the linguistic task requires a knowledge of the structure of language. This dimension is related to Cummins' context dimension in that contextualized uses of language usually do not require highly analyzed knowledge. The control dimension refers to the need to select and coordinate information. High cognitive control tasks require the selection and coordination of information. A task requires low cognitive control when the aspects relevant for solving the task are salient or when the subject has automatized the relevant control function. The control dimension is thus analogous to the cognitive-demandingness dimension proposed by Cummins, and is again the dimension of most concern here. Bialystok and Ryan (1985) state that "high control skills should transfer quite readily to other languages while tasks based on highly analyzed knowledge are likely to remain language specific. Thus, language learners should benefit most directly from previous language experience when operating in domains requiring high control."

In the metaphor task subjects are presented with novel sentences for which they must construct interpretations. Being novel, the task cannot be solved with purely automatized structures, and has been shown to require the use of mental effort, the amount of mental effort needed varying with the level of the interpretation. The task is thus a cognitively demanding (or a high cognitive control) one, and according to the above presented theoretical proposals, performance on such a task may be controlled more by general cognitive factors than by purely linguistic ones. In addition, one would expect to find linguistic interdependence in the context of such a task. Consistent with such proposals, the present study examines the extent to which a content-free (i.e., nonlinguistic) cognitive capacity constrains development of metaphor comprehension in L1 and L2. It also examines the extent to which the conceptual repertoire developed through L1 predicts level of metaphor comprehension in L2. The next section presents a possible cognitive-psychological interpretation of some of the theoretical constructs advanced above.

2. A COGNITIVE DEVELOPMENTAL PERSPECTIVE

The general theoretical framework used here is the theory of constructive operators of Pascual-Leone (1970, Pascual-Leone and Goodman 1979). This theory conceptualizes the psychological organism in terms of a modified "spreading activation" model composed of two interacting systems: (a) a content-specific system of schemes and (b) a content-free system of resource processes. Schemes are content-specific informational or knowledge-processing units. The subject's repertoire of schemes is his or her repertoire of knowledge structures. There are schemes for particulars of experience, for actions, for conceptual entities, for linguistic entities. The schemes one possesses are the result of one's learning history.

The only resource process to be discussed explicitly here is the M-operator. The M operator represents a limited amount of mental attentional energy (or mental capacity) that can be used to boost the activation of task-relevant schemes that are not sufficiently activated by the situation or by other organismic resources. In simple terms, M corresponds to the number of separate pieces of information, not directly activated by the input, that the subject can actively keep in mind at any one time. Such capacity is related to Spearman's (1927) g factor and to the modern notion of mental effort (Kahneman 1973) or working memory (e.g. Case, Marini, McKeough, Dennis and Goldberg 1986). This construct can be used to explicate, in part, the notions of "cognitively demanding" and "high cognitive control" tasks.

M is a limited-capacity resource. The maximal number of schemes that an individual can simultaneously boost with M is called his or her M-power. The theory posits that M power (when measured behaviourally) grows in a discontinuous fashion: Maximal M power increases linearly every other year during normal cognitive development from an M power of one mental unit at 3 and 4 years of age to an M power of 7 mental units at 15 years and older. Pascual-Leone (1970) has proposed that a growth in M power due to maturation is the transition rule in Piaget's developmental stages. There has been much empirical work supporting the validity of the proposed M power values — including research with various socioeconomic, cultural, and linguistic groups (Case 1975; DeAvila, Havassy and Pascual-Leone 1979; Globerson 1983; Miller 1980).

One can characterize a cognitively demanding task as one which requires the application of the M operator for its successful completion; this is opposed to a task whose solution is controlled by salient or overlearned aspects of the task situation. However, the degree of task demandingness is relative to the task and to the M power of the subject. One can characterize subjects in terms of their M power (by measuring it behaviourally). Similarly, one can characterize task performances terms of their M demand. The M demand of a task performance is the minimal amount of M power needed to generate the performance. This amount can be estimated by means of theory-guided task analysis. A task with an M demand of 3 mental units will normatively be demanding for a 7-year-old, less demanding for a 9-year-old, and beyond the capacity of a 5-year-old (the proposed M power of 5-, 7-, and 9-year-olds is, respectively, 2, 3, and 4 mental units). One source of the interdependence between L1 and L2 in cognitively demanding linguistic tasks may be the M power required for successful performance in either language.

M capacity is a content-free cognitive resource, an attentional energy that must apply on content structures or schemes. If the performance can be produced with overlearned or automatized structures, then little M capacity will be required for the performance (i.e., the task will be cognitive undemanding or require low cognitive control). If the task requires structures that are specific to a given language (e.g., a task requiring highly analyzed knowledge, such as the grammar of a specific language) then a main constraint on performance will be whether or not the child has learned those structures. If, however, the language task is a cognitively demanding one and it requires structures that are not specifically linguistic, then M power is likely to be the main constraint on performance. This is the claim made for the metaphor interpretation task, that it is a cognitively demanding task requiring experiential structures, which can be gained from interaction with the environment, and conceptual structures the learning of which may be facilitated by language learning, but which are not language specific (cf. Johnson, Fabian and Pascual-Leone 1986).

3. SUMMARY OF THE PRESENT STUDY AND HYPOTHESES

The study used a developmental design, with children of English and of Spanish home-background at each of three age levels: 7-8, 9-10, and 11-12 years. All children were assessed for M power, using a nonverbal measure, and within each age group the English annd Hispanic children were matched on age and M power score. The children were also tested for oral language proficiency and metaphor interpretation in English. The score for a metaphor item reflects the cognitive complexity of the metaphor interpretation and is scaled according to the M demand of the interpretation. The Hispanic children were also tested for oral language proficiency in Spanish, and a subsample were tested for metaphor interpretation in Spanish.

3:1 Hypotheses

- That there will be no overall difference between the Hispanic and English students in measured M power, although there will be overall differences in English oral language proficiency.

- That the developmental sequence of emergence of kinds of metaphor interpretations already found for monolingual English-Canadian children will be found for Hispanic children. That is, the same developmental scale of metaphor comprehension will apply across language groups and across languages.

- That across ages the metaphor performance of both Hispanic and English samples will be more strongly determined by age and by measured M power than by measured oral language proficiency.

- That relevant conceptual structures developed in L1 will facilitate metaphor comprehension in L2: For recent immigrants, oral language proficiency in Spanish will predict English metaphor performance better than does oral language proficiency in English.

- That on the metaphor task, Hispanic children with high proficiency in both Spanish and English will perform equal to or better than age-matched English monolinguals.

This last hypothesis is suggested by evidence that bilingualism yields cognitive benefits (e.g., Barik and Swain 1976, Ben-Zeev 1977, Bialystok and Ryan 1985, Cummins 1981, 1984a, DeAvila and Duncan 1980, Harley, Hart, and Lapkin in press, Ianco-Worrall 1972). If this is the case, one could expect bilingual students (at least those with high proficiency in both languages) to exhibit metaphor performance superior to that of age-matched monolingual students. In support of this hypothesis is Bountrogianni's (1984) finding that bilingual Canadian-Greek children performed better on a proverb interpretation task (in English) than did monolingual Canadian children. It is not suggested that bilingualism can increase one's M capacity (which is seen as increasing maturationally), but rather that it may foster growth of one's conceptual repertoire.

4 METHOD

4:1 Subject Selection

Subjects were initially recruited from grades 2, 4, and 6 of three separate schools in working-class areas of North Toronto. Parental permission forms were sent home with all children in these grades who, according to school records, had Spanish or English as the home language; the forms were written in the home language. Attached to the permission form was a questionnaire to be completed if parents agreed to their child's participation (see Appendix, p. 107). The questionnaire asked what had been the first language learned by the child and which languages were currently spoken (a) by the child at home, (b) by adults in the child's home, (c) by other children in the home, and (d) by the child outside the home. For each language they listed, parents rated on a five-point scale the frequency of use in the various situations.

All children with permission were tested in class groups with a nonverbal measure of mental capacity (i.e., M power). This measure was the Figural Intersections Test (FIT; de Ribaupierre and Pascual-Leone 1979, Pascual-Leone and Smith 1969); instructions were given in English and Spanish. Each FIT item shows a number of discrete geometric shapes on the right-hand side of a page and the same shapes in an intersecting configuration on the left-hand side. The number of shapes varies from two to eight; this number defines the class of an item. Subjects must locate in the intersecting configuration the one area where all shapes from the discrete set overlap. The FIT M-power score corresponds to the highest item class that the subject passes reliably (i.e., 75% correct), given reliable passing on all lower classes. Previous research has shown the FIT to be quite culture-free (DeAvila, Havassy and Pascual-Leone 1979, Miller 1980).

To ensure a normal developmental sample, subjects were selected to have an M-power score within the normal range for their age group. Specifically, subjects were selected to have an M-power score within one unit above or below the theoretically predicted, and usually found, M-power for their age (e.g., the predicted M-power of 7- and 8-year-olds is three mental units, thus, selected 7- and 8-year olds had a FIT score of of 3, 4, or 5). Subjects were selected into three age groups: 7-8, 9-10, and 11-12 years (henceforth these groups are referred to as 7, 9, and 11). Further criteria for selection of Hispanic subjects were that Spanish be their first language, their parents speak Spanish at home often or always, and the child speak Spanish at home at least sometimes. A further criterion for English children was that English be the only language spoken by the child and the parents; that is, the English sample was chosen to be monolingual.

Twenty Hispanic children per age group were selected, but because of the ethnic composition of the schools it was not possible to select a full matching sample of monolingual English children from the same schools. This necessitated going to a fourth school in a more middle-class area to obtain additional monolingual English children. Children with English home language in grades 2, 4, and 6 of the fourth school were tested with the FIT, and were added to the pool of English subjects. The final sample consisted of 20 Hispanic and 20 English children in each of three age groups; within age group the two language samples were matched on age and on M-power score. About 40% of the Hispanic children were immigrants from Latin America; the remainder were born in Canada. Table 1 (p. 91) shows descriptive statistics for the two samples (the "Standard Scores" and "Teacher Rating" are discussed below; note that the SD for age is in months). The language groups do not differ on age or measured M-capacity; M-capacity increases significantly with age, $F(2, 114) = 78.6$, $p < .001$.

4:2 Measurement of Oral Language Proficiency

The author (a native English speaker) tested the selected children individually in English with a metaphor interpretation task and an oral language proficiency task. Two to three weeks later a second female tester (a native Spanish speaker) tested all Hispanic children for oral language proficiency in Spanish and a subsample (10 subjects per age group) for metaphor comprehension in Spanish. Within language, testing for metaphor interpretation was followed by testing for oral language proficiency in a single half-hour session. The oral language proficiency test was the Oral Language subscale of the Woodcock Language Proficiency Battery (Woodcock 1980, 1981); English and Spanish versions of the test were used. The Oral Language subscale consists of three subtests: picture vocabulary, antonyms/synonyms, and verbal analogies (see Appendix for sample items). The content is not unlike that of a verbal IQ test. The test yields a number of scores: (a) a score for each of the subtests, (b) a summary score for the subscale (called the oral "cluster" score by Woodcock), and (c) a standard score which is like a deviation IQ (mean of 100 and SD of 15); scores (a) and (b) retain developmental variance, score (c) does not.

Descriptive statistics for the English and Spanish standard scores appear in Table 1. English children score significantly better on oral language proficiency (standard score) in English, $F(1, 114) = 63.97$, $p < .001$.

Immediately following the Woodcock, subjects were given a story retelling task (DeAvila and Duncan 1983). Children listened to a short tape-recorded story while looking at cartoon pictures illustrating the story. The child was then asked to retell the story in her own words. To date, the English story retelling data have not been scored. A native Spanish speaker scored the Spanish data on a 5-point level-of-oral-productive-proficiency scale (DeAvila and Duncan 1983).

In addition, the classroom teachers rated the English proficiency of their Hispanic students. This rating was done with the "Observation Form" from the *Language Assessment Scales* (DeAvila and Duncan 1983; see Appendix). Teachers rated on a 7-point scale (-3 to +3) their students' English proficiency in 10 areas. Items included ability to communicate in English in five different situations (e.g., explaining how to play a game, giving an oral report to the class) as well as ability in English pronunciation, comprehension, vocabulary, syntax, and general communicative competence. The overall score is the mean of the ratings on the 10 items.

4:3 Creation of Subsamples

Among the English children, subjects from the middle-class school scored significantly better on the Woodcock than did subjects from the working-class schools, F(1, 54, = 14.23, $p < .001$; the English sample was therefore divided into two subsamples, based on SES. Among the Hispanic children, recent immigrants scored significantly lower on the English Woodcock than did long-term residents, $F(1, 67) = 9.94$, $p = .002$, so the Hispanic sample was partitioned into two subsamples based on length of residence. Table 2 (p. 92) shows the mean scores of the subsamples on the independent variables; subsamples are middle- and working-class English, and Hispanic immigrants resident in Canada for three years or less versus Hispanics resident in Canada more than three years (most of these born in Canada; note that of the immigrants resident more than three years, all had in fact been resident at least five years).

The match on age and M power is maintained across the four subsamples (again M power increases significantly with age); but on the English standard score each subsample differs significantly ($p < .01$) from the others, with middle-class English scoring highest, followed by working-class English, long-term Hispanic residents, and finally recent Hispanic immigrants. On the Spanish standard score, the effect is reversed, with recent immigrants scoring significantly higher than long-term residents. Teacher rating of English proficiency was significantly lower for recent immigrants than for long-term residents, $F(1, 54) = 16.8$, $p < .001$. So on a nonverbal measure of developmental intelligence (i.e., M power) there are age differences, but no differences across subsamples. The subsamples clearly differ, however, on linguistic sophistication in English.

4:4 Dependent Variable: Metaphor Comprehension

The dependent variable was score on a metaphor interpretation task. Children were asked to interpret orally in English each of 11 ambiguous metaphors (see Table 3, p. 93). Data will be reported on six of the metaphors. These items have been used in previous research; the additional metaphors were practice or filler items. The metaphor items of interest were constructed by combining in a "______ was a ______" sentence-frame each of two subject terms (*my sister* and *my shirt*) with each of three predicate terms (*mirror*, *rock*, and *butterfly*). A subset of the Hispanic subjects were later tested for metaphor comprehension in Spanish. The Spanish metaphors of interest were direct translations of the six main English metaphors (see Table 3).

Subjects were first asked to define the nouns used in the metaphors, to ensure that they were familiar with the terms and their meanings. All subjects received the metaphors in the order listed in Table 3. Each metaphor item was read aloud, and the subject was asked to give a possible meaning for the sentence. Subjects sometimes gave more than one meaning for an item; they were given credit for their highest level spontaneous response to each item. The same procedure was followed for testing in English and Spanish; the subject was required to respond in the appropriate language. The metaphor-interpretation interviews were tape recorded for later transcription, and the transcripts were used in coding the interpretations. No identifying or other information on the subjects appeared on the transcripts; thus, the coding was done blind as to the subjects' age and language group (of course, language group was apparent in the Spanish metaphors).

Coding of metaphor interpretations. In terminology traditional to the field of metaphor, the subject term (e.g., my sister) is labeled the topic of the metaphor, and the predicate term (e.g., rock) the vehicle. The topic is what the metaphor is about, and the vehicle is the expression used metaphorically to say something about the topic. The author has proposed (Johnson 1982; Johnson, Fabian, and Pascual-Leone, in press; Johnson and Pascual-Leone 1984) that in comprehending a metaphor the subject selects some semantic aspect or facet of the vehicle (e.g., "physical hardness" for rock) and maps it to the topic. A "mapping" is some sort of mental transformation that the subject applies on the vehicle's semantics to accommodate it to the semantics of the topic. This model is novel in the metaphor literature, because it proposes different kinds of mapping processes which correspond to varying degrees of semantic transformation in metaphor comprehension.

It is proposed that the semantic process of comprehending a metaphor involves: (1) selecting some facets or aspects of the vehicle that are potentially applicable to the topic; and (2) then mapping these facets to the topic to evaluate analytically the appropriateness of the mapping. The mapping is done by means of semantic combinators; these are semantic transformations that convert one or more semantic facets into other different facets. The author posits various kinds of semantic combinators to represent varying degrees of semantic transformation. The term combinator serves to convey a mental operation that combines the semantics of the topic and of the vehicle to generate a metaphor interpretation. The author has developed a method for coding metaphor interpretations in terms of vehicle facets and semantic combinators. The combinators are developmentally ordered, based on the cognitive complexity (i.e., the M demand) of their application. Three semantic combinator kinds are discussed briefly here: the Identity, Analogy, and Predicate combinators. These are not the only combinators, but they are frequently used in metaphor interpretation, and they serve to convey the notion of increasing degrees of semantic transformation.

The Identity combinator is an instance of a low level of metaphor processing. In any metaphoric mapping, the subject first selects a facet or facets from his or her knowledge structure for the vehicle. In an Identity mapping the subject finds a facet in the vehicle that has (or could have) the same name and semantic definition in the topic and does a direct mapping of the facet from vehicle to topic. The facet is mapped without any change in meaning. An example is the following response to "My sister was a rock": "She was hard, like if you felt her hand you couldn't squish it or anything." The rock facet used is "hardness", the defining statement of which could be ((rocks do not change shape under the application of external physical force)). Here the subject selects a salient facet of rock and maps it to sister without changing the sense of the facet. Note that for a response to be scored as an Identity the mapped facet(s) must be compatible with the semantics of the topic; responses which violate this condition are coded with a lower level combinator.

A second example of an Identity mapping is the following response to "My sister was a mirror": "I could see myself in her eyes." The relevant vehicle facet is the mirror "image" facet, a description of which is ((a mirror gives back a reproduction or likeness in two dimensions of whatever is in front of it)). The facet is mapped unchanged from mirror to (a part of) sister; that is, the sister's eyes reflect back an image in the same way a mirror does.

The Analogy combinator is an instance of an intermediate level of metaphor processing. In an Analogy mapping the selected vehicle facets undergo a change in sense as they apply from vehicle to topic. This change represents an accommodation of the vehicle facets to the semantics of the topic. An example is the following response to "My sister was a rock": "She was unyielding; she had a hard personality." Here the "hardness" vehicle facet is accommodated to the topic by inserting topic-relevant content into the vehicle facet structure. So, referring back to the "hardness" rock facet, when it is applied to sister: "does not change shape" becomes "does not change behaviour" and "external physical force" becomes something like "verbal instruction" or "psychological pressure". In an Analogy mapping, the vehicle facet and the (semantically different) topic facet it maps are subsumed by a higher level and generic facet.

A second example of an Analogy mapping is the following response to "My sister was a mirror": "Maybe she was copying what other people do." This response is based on the mirror "image" facet described above, plus another facet referring to the movement of the image ((a mirror gives back an image that conforms to or reproduces changes in, or movements of, the object in front of it)). Here the facets are mapped to sister with a change in sense; that is, they are applied with the sense of behavioural rather than optical reproduction.

The Predicate combinator is an instance of the next higher level of processing; it applies within the topic after the vehicle-to-topic mapping. The Predicate serves to express the result of the initial mapping in terms that closely conform to the pragmatics of the topic. The Predicate is an elaboration of the mapped facet(s) by means of a concept or an instantiation that is relevant to the topic, but is not relevant to the vehicle. An example is the following response to "My sister was a rock": "She was mean." This response is based on an Analogical mapping of two rock facets, namely the "hardness" facet and a facet corresponding to the rock's capacity to hurt one. Over the result of this mapping the subject applies the Predicate or concept "mean"; thus someone who is psychologically "hard" and who can cause harm is described in sister terms as being "mean". Note that the concept "mean", which is based on the vehicle facets "hard" and "hurtful", cannot itself be applied to rock (rocks may be hard, but they're not mean!); it can only apply to the topic (sister).

Another Predicate example is the following response to "My sister was a mirror": "Like the sister is like a mirror -- like the sister would usually do most the same stuff that you do. Like would copy what you, like just say my sister, like I buy a chocolate bar and then she'd go 'Mom, can I buy a chocolate bar? Cause she bought one.' Like she'd do what the other person did." This response is based on an Analogical mapping like the one described above ("she was copoying"), but now the subject instantiates (Predicate) this Analogy by describing a particular example of copying behaviour.

Metaphor score. These three semantic combinators are progressive accommodations of the vehicle facets to the semantics of the topic. Using Pascual-

Leone's method of task analysis (e.g., Pascual-Leone 1980; Pascual-Leone and Goodman 1979), the mental-capacity demand (i.e., the cognitive-developmental complexity) of these and other semantic combinators has been estimated, and these complexity estimates have been used to construct a quantitative scale of metaphoric processing (Johnson 1982; Johnson and Pascual-Leone 1984). Note that level of metaphoric processing is defined in terms of degree of accommodation of the vehicle facets to the semantics of the topic. M demand actually increases with processing level, because each successive level requires the subject to take into consideration more semantic aspects of the topic.

The quantitative metaphor-interpretation scale ranges from 0 for complete failure to make a semantic mapping (e.g., denial that the sentence is meaningful) to 7 for a conceptual Predicate (the instantiation type of Predicate is scored slightly lower). The validity and reliability of this scale have been established in previous work with English monolingual children (Johnson 1982, Johnson and Pascual-Leone 1984). In this earlier work, metaphor score increased significantly with age and, across ages, was highly correlated with chronological age, $r(160) = .80$, mental capacity, $r = .67$, and mental age, $r = .77$. The just described method was used to code and quantify the metaphor interpretations in the present study. On the 8-point coding scale there was 81% agreement between the author and an independent coder who scored a subset of the transcripts.

5. RESULTS: METAPHORS INTERPRETED IN ENGLISH

5:1 Full English and Hispanic Samples

Figure 1 (p. 101) illustrates results for the full English and Hispanic samples on the English metaphors. The top curves (A) show the relation between age, language group, and the quantified metaphor score averaged across the six English metaphor items. Performance increases with age, and the two curves are colinear. The colinearity suggests that the age-bound causal factor in metaphor comprehension is largely independent from the specific language-bound causal factor. English subjects perform somewhat better than Hispanic subjects.

An Age x Language-group x Topic x Vehicle (3 x 2 x 2 x 3) analysis of variance was performed on the quantified metaphor score. There was a large and significant main effect for Age group, $F(2, 113) = 87.71$, $p < .001$, as well as a significant effect for Language group, $F(1, 113) = 9.01$, $p < .003$.

The lower curves in Figure 1 (B) illustrate a significant Age x Language-group x Topic effect, $F(2, 113) = 3.64$, $p = .029$; at least at the older ages, the Language-group effect seems to be due mainly to lower Hispanic performance on items with the shirt topic. This effect may be due to a deficient linguistic repertoire for discussing shirts in the Hispanic children. These children speak Spanish at home, and it is the home context in which shirts are most likely to be discussed. Away from home, the children speak mainly English, and characteristics of sisters (and of other people) seem more likely to be topics of conversation in this context than are properties of shirts.

Additional significant effects were for (a) Age x Topic, $F(2, 113) = 5.1$, $p = .007$; (b) Vehicle, $F(2, 226) = 33.2$, $p < .001$; and (c) Topic x Vehicle, $F(2, 226) = 3.6$, $p = .029$. (a) Seven-year-olds tended to do better on items with the shirt topic, whereas older subjects did better on sister-topic items. (b) Overall, the best performance was on items with the rock vehicle; (c) this vehicle effect was strongest for items with the shirt topic.

Results confirm, as predicted, that the same developmental scale of metaphor comprehension applies to both language groups. An analysis of covariance was done to determine whether differences in English oral language proficiency would account for the main effect of language group; English standard score was used as covariate in an Age x Language-group x Topic X Vehicle analysis. This analysis yielded a significant main effect for Age, $F(2, 112) = 90.9$, $p < .001$; but no significant effect for Language-group. The significant Age x Topic, Age x Language-group x Topic, Vehicle, and Topic x Vehicle effects described above were obtained again. Thus controlling for English proficiency eliminates the main effect for Language, but not the interaction.

Figure 2 (p. 102) illustrates results when the subjects are partitioned by their measured M-power (i.e., score on the FIT) rather than by age. For both language groups, performance increases with increasing M-power, and again the curves are colinear. The greatest performance increase is across M-power levels of 3 to 5 mental units; this range corresponds to the estimated range of M-demands of the proposed levels of metaphoric processing (Johnson 1982; Johnson and Pascual-Leone 1984). An M-power x Language-group x Topic x Vehicle (3 x 2 x 2 x 3) ANOVA was performed on the metaphor score. In this analysis subjects with measured M-power of 2 and 3 were combined into one group, and subjects with M-power 5 and 6 were also combined; this was done to avoid small cell sizes and seemed justified due to small metaphor performance differences at the extremes of the empirical M-power scale. Thus the M-power groups in the ANOVA were M power of 2-3, 4, and 5-6 mental units. There was a significant effect for M-power, $F(2,113) = 49.9$, $p < .001$, as well the various effects reported above, involving language-group, topic, and vehicle. There were significant interactions for M-power x Topic and M-power x Language-group x Topic, paralleling those found with Age. English subjects showed no effect of Topic. Among Hispanic subjects, those in the lowest M-power group performed better with shirt items than with sister items; those in the highest M-power group performed better with sister; and those with M-power of 4 mental units showed no topic effect. In the highest M-power group, the language effect was due to lower Hispanic performance on shirt items; in this group performance on sister items was comparable across Hispanic and English subjects. The tendency for younger (and lower M power) subjects to perform better with shirt items and for older (and higher M power) subjects to perform better with sister items has been found in previous research (Johnson and Pascual-Leone 1984; Johnson, Fabian, and Pascual-Leone, in press).

5:2 Subsamples

Analyses of variance. Recall that due to differences in performance on the English oral language proficiency test the two language groups were divided into subsamples: middle-class vs. working-class English, and recent immigrant vs. long-term resident Hispanics. Figure 3 (p. 103) shows the relation between age and mean metaphor score for the four subsamples. Again, the curves are colinear and all subsamples show a strong and significant performance increase with age.

A comparison of the two English subsamples yielded a significant main effect for social class, $F(1, 54) = 5.13$, $p = .028$; children from working-class areas perform significantly less well on metaphor than do children from a middle-class area. The English subsamples also exhibit the main effect for vehicle described above, $F(2, 108) = 15.0$, $p < .001$.

Comparing the two Hispanic subsamples, there is no significant effect of length of residence on metaphor score, $F(1, 53) < 1.0$. The Hispanic subsamples exhibit a main effect for topic, $F(1, 53) = 6.7$, $p = .013$; overall, performance on items with the sister topic is better than that on items with the shirt topic. Hispanics also show the significant effects of Topic x Age, $F(2, 53) = 7.3$, $p < .002$; Vehicle, $F(2, 106) = 16.4$, $p < .001$; and Topic x Vehicle, $F(2, 106) = 5.7$, $p < .005$, described above.

Although Hispanic children clearly perform lower than middle-class English children, $F(1, 79) = 15.7$, $p < .001$; when Hispanic children are compared with their working-class schoolmates there is no main effect for language-group, $F(1, 87) = 1.66$, $p > .20$. The Hispanic vs. English working-class comparison yields a significant Topic x Language-group effect, $F(1, 87) = 6.52$, $p < .012$; the two language groups perform at the same level on items with the sister topic, but English subjects perform better on shirt-topic items. The Hispanic versus English-subsamples comparisons yield additional Topic and Vehicle main and interaction effects, which are not of great interest here.

Figure 4 (p. 104) shows results for the subsamples plotted by measured M-power rather than by age. All subsamples show a significant increase in metaphor score with increasing M-power. ANOVAs using M-power as a factor, rather than Age, yielded results comparable with those reported above.

It would appear that, overall, socioeconomic class has a greater influence on level of metaphor interpretation than does language group. This may be due to the opportunities to enrich one's conceptual repertoire (in addition to the specifically linguistic repertoire) that a middle-class upbringing provides. Despite having significantly lower English proficiency, Hispanic subjects do not perform worse on metaphor than do their working-class English schoolmates. This is consistent with the claim that the metaphor task is relatively insensitive to specifically linguistic variance. It appears rather to measure the conceptual level and the constructive complexity (i.e., the cognitive-developmental difficulty) of metaphor interpretaions. As predicted, age and M power exert the major effect on level of metaphor interpretation.

Correlations. Table 4 (p. 94) contains across ages correlations for the English subsamples (correlations for the full English and Hispanic samples appear in the Appendix, p. 107). The metaphor score is the mean across the six items. The M-power score is performance on the nonverbal Figural Intersections Test. There are five Woodcock oral-language proficiency scores (prefaced with "E-" to denote English proficiency): a score for each of three subtests (i.e., picture vocabulary, antonyms/synonyms, and verbal analogies), a summary score (the aforementioned scores retain developmental variance), and a standard score (normalized for age). The following discussion focuses on the correlations of metaphor with age, M power, and English-summary.

Metaphor is highly correlated with age in both subsamples. For middle-class subjects, the correlations of age and M power with metaphor are higher than the correlation of linguistic proficiency with metaphor. For working-class subjects, the correlations of age and linguistic proficiency with metaphor are higher than the correlation of M power with metaphor. In native speakers, when linguistic repertoire is strong (i.e., middle class), relatively more developmental variance in metaphor is accounted for by general cognitive capacity than by specific linguistic skill. When linguistic repertoire is weaker (i.e., working class), relatively more variance in metaphor is accounted for by differences in linguistic skill.

Table 5 (p. 95) contains across ages correlations for the Hispanic subsamples. Additional scores appearing in this table are those for Spanish (i.e., "S-") Woodcock oral-language proficiency, Spanish story-retelling, and teacher rating of proficiency in English. The focus again is on correlations of metaphor with the age, M power, and summary-language scores. The correlations of age and M power with metaphor score are high in both subsamples (recall that M-power score is derived from a nonverbal measure). For Hispanic children resident more than three years, the correlation of metaphor with English proficiency is higher than the correlation of metaphor with Spanish proficiency. For children resident three years or less, the correlation with Spanish proficiency is higher than that with English; this is evidence for the prediction that conceptual structures developed in the first language will facilitate metaphor comprehension in the second language.

Let us examine the correlations between English and Spanish oral language proficiency. For children resident more than three years, the correlations are fairly high. Given a reasonable level of proficiency in English, development in one language is accompanied by development in the other. For children resident three years or less the correlations are low and nonsignificant, with one exception. The exception is the verbal analogies subtest. This subtest had fairly simple vocabulary and seemed to tap the ability to find relations using language, that is, it appeared to be more a conceptual than a specifically linguistic task. For subjects with very limited English proficiency, there is evidence for linguistic interdependence when the task emphasizes conceptual ability rather than linguistic skill. It would appear that conceptual structures developed by means of the first language do transfer to the second language, but such transfer is exhibited only when the second-language task does not penalize too heavily for limited proficiency; in the current study, this appears to be the case for the metaphor and verbal analogies tasks.

Regression analyses. Stepwise multiple regression analyses are used to summarize and clarify the sources of variance in metaphor development in the four subsamples. Results of the stepwise regressions are shown in Table 6 (p. 96). The dependent variable was scored on the metaphor task; the independent variables were entered in the following order: M power, English proficiency, and chronological age. This order should serve to purify the sources of variance. M power would appear to be the purest source of variance — a nonverbal measure of developmental intelligence. Entered after M power, the English proficiency measure should retain variance that is more specifically linguistic. Finally, age entered last should carry variance reflective of general experience. The table gives the cumulative variance accounted for (i.e., R^2) at each step in the regression, and the R^2-change as each new variable is entered.

For English working-class children, language proficiency contributes a good deal of variance to metaphor score, even after M power has been partialled out; age adds some variance beyond language. For English middle-class subjects, there is little variance unique to language proficiency once M power has been partialled out; age contributes variance beyond that carried by M power. For recent Hispanic immigrants, M power is the major source of variance, English proficiency contributes no unique variance, and age contributes some variance beyond M capacity. For longer-term Hispanic residents, M power is again the major source of variance, but English proficiency and age each contribute some additional variance. These results are in line with the prediction that, across ages, metaphor performance would be more strongly determined by age and by measured mental capacity than by measured oral language proficiency. Note that for Hispanic subjects when Spanish proficiency is entered before English proficiency, Spanish contributes some unique variance to metaphor performance in English.

If the linguistic environment is rich (as it is likely to be in the middle class), then mental capacity may set the limit on performance level in metaphor interpretation. Mental capacity needs structures on which to apply, however, and if the linguistic environment is less rich (as it is likely to be in the working class), then linguistic repertoire is likely to play a greater role in constraining metaphor comprehension. An interesting result is that in the regression analyses, Hispanic subjects seem to exhibit the correlational pattern of English middle-class children, despite the low English proficiency scores of the Hispanic children. A major cognitive benefit of bilingualism may be its ability to develop in children an enriched repertoire of conceptual/logical structures. I use "conceptual/logical" to refer to language-based, but not specifically linguistic, logical structures. This repertoire is not reflected in the measured English proficiency of the Hispanic subjects but is reflected in their metaphor performance.

5:3 Analysis of Proficient Bilinguals

It was hypothesized that, on the metaphor task, Hispanic children with high proficiency in both Spanish and English would perform equal to or better than age-matched English monolinguals. To test this prediction a different subsample partitioning of the subjects was made. Within each age group of the Hispanic sample, six "proficient bilingual" subjects were selected, that is, subjects showing evidence of good proficiency in both languages. In practice, good proficiency in English was an English-Woodcock standard score of 83 or higher. Good proficiency in Spanish was a Spanish-Woodcock standard score of 82 or higher _or_ a standard score in the 70's accompanied by a Spanish story-retelling score at the "proficient speaker" level (i.e., a score of 4 or 5). These were the most stringent criteria that could be applied and still yield a minimally acceptable sample size.

The "proficient bilingual" (henceforth called "bilingual") sample was used in a number of analyses. In a first analysis they were compared with the English subjects who had been matched with the bilinguals in the initial subject selection (i.e., matched on age and _M_-power score). Age x Language-group, 3 x 2, ANOVAS were done on _M_-power score, English standard score, and metaphor score. None of the analyses yielded a significant (main or interaction) effect for Language-group. Thus, the weak version of the hypothesis -- that proficient bilinguals would perform as well as age-matched monolinguals -- was confirmed. There was no evidence, however, for bilingual superiority in metaphor interpretation.

Similar analyses were done comparing the bilinguals with English subjects who were selected to match the bilinguals, within age group, on English standard score. Again, there were no significant effects for Language-group.

To determine whether the good performance of the Spanish bilinguals, relative to matched English subjects, was due to their bilingualism or due simply to their being selected to have high English scores, the bilinguals were compared with other Hispanic subjects. Within age group, each proficient bilingual subject was matched with an Hispanic subject whose English standard score was about the same, but whose Spanish standard score was low. One might characterize the second sample as "marginal bilinguals". Age x Bilingual-group ANOVAS yielded a significant Bilingual-group effect for Spanish standard score (proficient bilinguals scored higher), but no other effects for degree of bilingualism.

Larger sample sizes would likely be needed to provide a definitive answer, but the analyses above suggest that, with the sort of measures used here, bilingualism yields neither benefits nor deficits in level of metaphor interpretation.

6. RESULTS: METAPHOR INTERPRETATIONS IN SPANISH

6:1 Sample and Procedure

As mentioned above a subsample of the Hispanic subjects was tested for metaphor comprehension in Spanish. The subsample consisted of 10 subjects per age group, selected to be representative of the full age-group sample on the English Woodcock standard-score. "Representative" was defined in terms of the subsample's having about the same mean score and dispersion of scores as the full sample. Equal numbers of males and females were selected. Table 7 (p. 97) lists the mean scores by age group for the selected Hispanic subsample. By comparing with the full Hispanic-sample means listed in Table 1 (p. 91), the reader can ascertain that the subsample is highly representative of the full sample on the independent variables and on metaphor performance in English.

The Spanish metaphor interpretations were coded by the author in close consultation with a native Spanish speaker, the person who had done the Spanish testing; the latter was familiar with the metaphor coding system.

6:2 Analyses of Variance

Figure 5 (p. 105) illustrates the metaphor performance in English and Spanish of the subsample. In both languages performance increases with age, although in Spanish the difference between ages 9-10 and 11-12 is diminished. In the youngest age group, metaphor performance is better in Spanish than in English; at 9- to 10-years, performance is about the same in the two languages; and in the oldest age group English performance is better.

A 3 x 2 x 2 x 3, Age x Language x Topic x Vehicle ANOVA, with repeated measures on the last three factors ("Language" refers to English vs. Spanish metaphors), was performed on the metaphor scores. The analysis yielded significant effects for Age, $\underline{F}(2,22) = 18.2$, $\underline{p} < .001$; Age x Language, $\underline{F}(2,22) = 4.75$ $\underline{p} < .02$; Vehicle, $\underline{F}(2, 44) = 23.9$, $\underline{p} < .001$; and Topic x Vehicle, $\underline{F}(2, 44) = 9.9$ $\underline{p} < .001$.

The Age and Age x Language effects were discussed briefly in the context of Figure 5 above. The better Spanish metaphor performance of the 7-8-year-olds was caused by an increase in number of interpretations at the Identity level. The better performance in Spanish of the 9-10-year-olds was caused by an increase in interpretations at the Analogy level. The worse performance in Spanish of the 11-12-year-olds was caused by a decrease in interpretations at the Predicate level. The performance differential between the two languages was due to better or worse performance at the metaphor processing level predicted to be within the $\underline{M}$ capacity of the age group (according to results of task analyses, 7-8 year olds have the $\underline{M}$ capacity for Identity-level responsing, 9-10's the capacity for Analogy-level responding, and 11-

12's for Predicate-level; cf. Johnson 1982, Johnson et al. in press, Johnson and Pascual-Leone 1984).

Items with the *rock* vehicle again produced the highest scores. There was a Topic effect (*sister* items easier than *shirt* items) only for items with the *butterfly* vehicle. For *sister* items, the *rock* vehicle was easiest, closely followed by *butterfly*, and then *mirror*. For *shirt* items, the order was *rock* easiest, then *mirror*, then *butterfly*. These same Topic x Vehicle effects have been found in previous research (Johnson 1982).

Results confirm, as predicted, that the same developmental scale of metaphor comprehension applies across languages (at least for English and Spanish). The specific reasons for the language differentials in metaphor performance are not immediately apparent. The performance increase from English to Spanish (7- to 10-year-olds) may be due to being tested in the first language or may simply be a practice effect, subjects having been tested first in English. (Note, however, that the increases occur only within the metaphor processing level appropriate for the age group.) The performance decrease from English to Spanish at 11-12-years could be due to an insufficient linguistic repertoire for responding at the Predicate level in Spanish. However, the decrease in Predicate interpretations at this age is accompanied not by an increase in responding at the Analogy level, but by an increase at the Identity level. This suggests that the reason may be motivational. The tester noted that the oldest subjects seemed more reluctant than the other ages to be tested in Spanish; as well these subjects may have found it uninteresting to be tested a second time on basically the same task. This disinterest could have manifested itself in an increased tendency to give minimally effortful and minimally acceptable responses (i.e., Identities). Testing of a larger sample with order of language counterbalanced could provide clearer answers.

A 2 x 2, Length-of-Residence x Metaphor-Language ANOVA yielded no significant effects for Length-of-Residence ($\leq$ 3 yrs. vs. $>$ 3 yrs.). Length of residence appears to have no effect on the metaphor-performance level of Hispanic subjects in either language.

A 3 x 2 x 2 x 3, *M*-power x Language x Topic x Vehicle ANOVA, with repeated measures on the last three factors was performed on the metaphor scores; as in the previous analyses, the three *M*-power levels corresponded to scores of 2-3, 4, and 5-6 on the FIT. The analysis yielded significant effects for *M*-power, $F(2,22) = 15.2$, $p < .001$; *M*-power x Language, $F(2,22) = 3.98$, $p = .03$; *M*-power x Topic, $F(2, 22) = 5.45$, $p = .01$; *M*-power x Language x Topic, $F(2, 22) = 3.6$, $p = .04$; as well as the Vehicle and Topic x Vehicle effects reported above.

The *M*-power x Language x Topic effect is illustrated in Figure 6 (p. xxx). At the lowest *M*-power level, subjects perform better in Spanish with both topics. At the middle *M* level performance in Spanish is superior only on items with the *shirt* topic. At the highest *M* level there is a cross-over: Subjects perform somewhat better in English within each topic; but across languages, performance on *sister* items is better than that on *shirt* items. Performance increases with increasing *M* power in both languages and with both topics; only with Spanish performance on *shirt* is there little increase between the middle and highest *M* group. The reason for this could be the motivational one mentioned above, a tendency to produce the least effortful acceptable response. There are more Identity-level responses that are minimally acceptable for *shirt* than for *sister*, and the flattening of the Spanish *shirt* curve is largely due to a tendency for the high *M* subjects to respond with an Identity (i.e., "My shirt was colourful") to the item "My shirt was a butterfly". The tendency for subjects with M power of 5 mental units to respond with Identities on this item has been found also with English subjects in previous research (Johnson 1982).

6:3 Correlations

Table 8 (p. 98) contains across ages correlations for the subjects tested in Spanish. The correlations between the English metaphor score and the independent variables are almost the same for the subsample as they were for the full Hispanic sample (as the reader can ascertain by comparing with the table of correlations in the Appendix, p. 113). The correlation between the English and Spanish metaphor scores is sizeable (r = .81); the correlations of Spanish metaphor with the independent variables are somewhat lower than those of English metaphor. Both metaphor scores correlate highly with age, and for both scores the correlation with M-power is at least as high as the correlations with the Woodcock language scores. English metaphor score correlates more highly with English Woodcock, whereas Spanish metaphor correlates more highly with Spanish Woodcock. The correlations further support the claim that metaphor interpretation is constrained more by developmental factors than by specifically linguistic factors.

Table 9 (p. 99) contains correlations of the English and Spanish metaphor scores with selected independent variables, when the subsample is partitioned into recent immigrants (resident 3 years or less, n = 13) versus long-term residents (more than 3 years, n = 17). The n's are rather low here, so the discussion focuses on the patterns of correlation rather than on magnitudes. One may assume that for the recent immigrants Spanish is the dominant language (their mean Woodcock English standard score was 75.3, their Spanish standard score 89.2), and for the long-term residents English is the dominant language (mean English standard score 90.8, Spanish standard score 77.3). The essence of the correlations in Table 9 may be that for the non-dominant language, developmental factors (age, M-power) predominate in accounting for developmental variance in metaphor interpretation. In the dominant language linguistic proficiency takes on increased importance. These patterns are further explored in the regression analyses reported below.

6:4 Regression Analyses

Again stepwise multiple regression analyses were used to summarize and clarify the sources of variance in metaphor development. Table 10 (p. 100) contains the results of various regression analyses in terms of cumulative R^2 and R^2-change at each step. The purpose of these regressions was not to obtain predictive regression equations, but to gain understanding of the sources of variance. The regressions on the full subsample tested in English and Spanish reveal that mental capacity (i.e., M power, when entered first in the regression) is the main source of developmental variance in metaphor interpretation in both languages. Age, entered last, also contributes variance in both languages. Language-based variance (Woodcock summary scores) is somewhat higher for the English than for the Spanish metaphors.

Table 10 also contains stepwise regressions for the subsample partitioned by length of residence; again the focus is on the regression patterns rather than the magnitudes, because the n's are low here. Again M-power score, when entered first, is the main source of variance in all scores except one: For recent immigrants Spanish metaphor performance is mainly constrained by proficiency on the Spanish Woodcock. Note that Spanish Woodcock score also contributes a fair bit of variance to the English metaphor performance of these subjects.

The regressions by length of residence, on the English metaphor score, have essentially the same pattern as those discussed previously for the full Hispanic sample partitioned by length of residence (see Table 6, p. 96). The regressions by length of residence, on the Spanish metaphor score, are similar in pattern to the regressions by social class for the English sample (see Table 6; in Table 10, reference is to the regressions using Spanish summary score as independent variable). Recall that for the middle-class English children, M power was the major source of variance in metaphor performance; whereas, for working-class English children language proficiency was also a main source of variance. The interpretation was made that if the linguistic environment is rich then mental capacity may set the limit on performance level in metaphor, but if the linguistic environment is less rich then linguistic repertoire is likely to play a greater role in constraining metaphor comprehension. It was also suggested that a cognitive benefit of bilingualism maybe its ability to develop in children an enriched repertoire of conceptual/logical structures. It may be the case that vis-à-vis the native language the recent immigrants' exposure to English may not have been sufficient to have had this enriching effect, and these subjects thus reveal a correlational pattern similar to that of the English working-class children.

7. CONCLUSIONS

Let us first evaluate the hypotheses. The subject samples (and subsamples) were matched on age and on a nonverbal measure of mental capacity, but they differed in measured oral language proficiency. It was proposed that the metaphor task, although clearly a language task, measured mainly level of conceptualization in metaphor comprehension, and thus performance would be constrained more by cognitive-developmental than by specifically linguistic capacities. This hypothesis was clearly upheld: In all analyses of variance the size of the main effect for Age (or for M power) greatly outweighed any sample effect (e.g., SES group, Language group). The metaphor measure, the coding scheme and quantified score, was further validated; it captured well development in metaphor interpretation, in both English and Spanish, in Hispanic children and, in English, across two SES levels of monolingual English children.

It was proposed that the metaphor task — being a cognitively demanding conceptual task — would show evidence of a "common underlying proficiency" (Cummins et al., 1984) in the Hispanic children. The author suggests that an important aspect of interdependence across languages in cognitively demanding tasks is in fact not linguistic, but is the developmental mental capacity (M power) needed to solve the task. Beyond this, however, there was evidence for interdependence more clearly related to language. This was in the correlations in the Hispanic children between Woodcock English proficiency and Spanish metaphor performance and between Woodcock Spanish proficiency and English metaphor. When the metaphor testing was in the "non-dominant" language (i.e., English for recent Hispanic immigrants, Spanish for long-term Hispanic residents) the cross-language correlation (between Woodcock and metaphor) was generally higher than the within-language correlation. The basis of these correlations is likely conceptual structures that are developed through language, but are not specifically linguistic. These findings support the position of interdependence between L1 and L2, at least at the conceptual level.

Note that the Woodcock itself measures a particular type of proficiency; its content — vocabulary, antonyms/synonyms, analogies — is also fairly conceptual, although it requires finding the right words and so requires some specifically linguistic knowledge. Other measures of proficiency in the present study (the teacher rating of English proficiency and the Spanish story re-telling task) appear to be less conceptual. These measures seem to be valid, judging from their correlations with the Woodcock, but they did not exhibit high correlations with metaphor performance.

It was hypothesized that Hispanic children with good proficiency in both English and Spanish would perform as well as or better than age-matched English monolinguals. The weak version of this hypothesis (i.e., equal performance) was confirmed, but there was no evidence of "bilingual" superiority in level of metaphor performance.

The study does, however, contain evidence for possible cognitive benefits of bilingualism. This is in the fact that despite very low scores on Woodcock English proficiency, the Hispanic children do no worse on metaphor than do their working-class schoolmates; and in the fact that the sources of variance in the Hispanic children's English metaphor performance are similar to those of the middle-class English children (see the regression analyses in Table 6, p. 96). Again it is suggested that a benefit of bilingualism may be the development of an enriched repertoire of conceptual structures. This enrichment is not evidenced in the Woodcock proficiency scores of the Hispanic children (perhaps because this test demands too much specifically linguistic knowledge, but it is evidenced in their good English metaphor performance and in the fact that this performance is more constrained by developmental variance (i.e., M power) than by language (Woodcock) variance.

Most of the Topic and Vehicle effects in the present study have been found as well in previous research (Johnson 1982, Johnson and Pascual-Leone 1984). The developmental increase in the overall metaphor score (averaged across the six items) seems due to increase in mental capacity. The topic and vehicle effects, however, may be due more to effects of learning or motivation. Consistently, the best performance is on items with the rock vehicle. At all ages the most frequently used rock facet is "hard". Sister-relevant senses of "hard" are used commonly in everyday speech (e.g., "hard on me", "hard personality", "hard to get along with" — all of these coded as Analogies), and some subjects may have previously learned these senses. As well there are a number of Predicates based on "hard" that may have been overlearned by the children (e.g., for shirt — "starchy" or experiential situations were a shirt was made hard — laundered badly, frozen out on the clothesline; for sister — "mean" or "stubborn").

Young children tend to perform better with shirt items and older children with sister items. This is likely due to an interaction between the item type and the processing capacity of the subject. Shirt shares more physical facets with the vehicles than does sister. This means that Identity mappings are more likely to be compatible with the semantics of shirt than with those of sister. Since the youngest children have only the capacity for Identity-level responding, they have more chance to exhibit this capacity with the shirt items. An adequate interpretation of a sister-topic item is more likely to require processing at the Analogy or Predicate level; the older children have the capacity for responding at these higher levels, and they perform better with sister than with shirt.

At least at the older age levels, the lower performance of Hispanic children relative to English children is due to lower Hispanic performance on shirt items. Their performance on sister items is comparable to that of English subjects (see Figure 1B, p. 101). As discussed earlier, this effect could be linguistic -- that the Hispanic children do not have the linguistic repertoire to express themselves at a high level on the shirt items (although their vocabulary for talking about sisters is adequate). The effect could also be motivational, shirts are less interesting topics of conversation than are sisters, and for shirt an Identity is a minimally appropriate (and minimally effortful) response; for sister Identities are less acceptable. I speak here of acceptability to the subject; the interviewer accepts all interpretations without comment.

In conclusion, the study demonstrates the validity of the metaphor-interpretation measure, across languages and across language groups. The metaphor task is a cognitively demanding conceptual task, which appears to be more sensitive to developmental than to linguistic variance. Despite lower scores on a standardized English proficiency measure, Hispanic children do not perform more poorly on English metaphor than do their working-class English schoolmates, and Hispanic children show the same correlational pattern as do middle-class English children. Metaphor is of the essence of language, yet appropriately measured, metaphoric proficiency is related more to general developmental level than to specific linguistic proficiency. This would seem to indicate the advantage of measuring conceptual/logical intelligence by means of metaphor processing instruments rather than by the so-called verbal IQ tests. This change might render the assessment of conceptual/logical intelligence less prone to be discriminatory in a multicultural society like ours.

Footnotes

[1] Preparation of this chapter was assisted by a York University Faculty of Arts Minor Research grant to the author. The author acknowledges the advice of Drs. E. De Avila and J. Pascual-Leone.

References

Barik, H.C., and Swain, M. (1976). A longitudinal study of bilingual and cognitive development. **International Journal of Psychology**, 11, 251-263.

Ben-Zeev, S. (1977). Mechanisms by which childhood bilingualism affects understanding of language and cognitive structures. In P.A. Hornby (Ed.), **Bilingualism: Psychological, social, and educational implications**. New York: Academic Press.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. **Child Development**, 57, 498-510.

Bialystok, E., and Ryan, E. (1985). A metacognitive framework for the development of first and second language skills. In D.L. Forrest-Pressley, G.E. MacKinnon, and T.G. Waller (Eds.), **Meta-cognition, cognition, and human performance**. New York: Academic Press.

Bountrogianni, M. (1984). **A cross-cultural study of children's metaphoric capacity.** Paper presented at the meeting of the American Psychological Association, Toronto.

Case, R. (1975). Social-class differences in intellectual development: A neo-Piagetian investigation. **Canadian Journal of Behavioural Science**, 7, 244-261.

Case, R., Marini, Z., McKeough, A, Dennis, S., and Goldbert, J. (1986). Horizontal structure in middle childhood: Cross-domain parallels in the course of cognitive growth. In I. Levin (Ed.), **Stage and structure: Reopening the debate**. Norwood, N.J.: Ablex.

Cummins, J. (1981). The role of the primary language in promoting educational success for language minority students. In California State Department of Education (Ed.), **Schooling and language minority students: A theoretical framework**. Los Angeles: National Dissemination and Assessment Center.

Cummins, J. (1984a). Bilingualism and cognitive functioning. In S. Shapson and V. D'Oyley (Eds.), **Bilingual and multicultural education: Canadian perspectives**. Clevedon, Avon, England: Multilingual Matters.

Cummins, J. (1984b). Wanted: A theoretical framework for relating language proficiency to academic achievement among bilingual students. In C. Rivera (Ed.), **Language proficiency and academic achievement**. Clevedon, Avon, England: Multilingual Matters.

Cummins, J., Swain, M., Nakajima, K., Handscombe, J., Green, D., and Tran, C. (1984). Linguistic interdependence among Japanese and Vietnamese immigrant students. In C. Rivera, (Ed.), **Communicative competence approaches to language proficiency assessment: Research and application**. Clevedon, Avon, England: Multilingual Matters.

DeAvila, E., and Duncan, S. (1980). **The language minority child: A psychological, linguistic, and educational analysis** (Research Report). Larkspur, Ca.: DeAvila, Duncan and Associates.

DeAvila, E., and Duncan, S. (1983). **Language Assessment Scales (LAS).** San Rafael, Ca.: Ca.: Linguametrics Group.

DeAvila, E., Havassy, B., and Pascual-Leone, J. (1976). **Mexican-American schoolchildren: A neo-Piagetian approach.** Washington, D.C.: Georgetown University Press.

de Ribaupierre, A., and Pascual-Leone, J. (1979). Formal operations and M power: A neo-Piagetian investigation. **New Directions for Child Development, 5,** 1-43.

Globerson, T. (1983). Mental capacity and cognitive functioning: Developmental and social class differences. **Developmental Psychology, 19,** 225-230.

Harley, B., Hart, D., and Lapkin, S. (in press). The effect of early bilingual schooling on first language skills. **Applied Psycholinguistics, 7.**

Ianco-Worrell, A. (1972). Bilingualism and cognitive development. **Child Development, 43,** 1390-1400.

Johnson, J. (1982). The development of metaphor comprehension: Its mental-demand measurement and its process-analytical models. Unpublished doctoral dissertation, York University, Toronto.

Johnson, J., Fabian, V., and Pascual-Leone, J. (1986). Quantitative hardware-stages that constrain language development. Manuscript submitted for publication.

Johnson, J., & Pascual-Leone, J. (1984). **Level of processing and mental-attentional demand in metaphor comprehension: Their measurement and developmental validation** (Research Rep. No. 149). Toronto: York University, Department of Psychology.

Kahneman, D. (1973). **Attention and effort.** Englewood Cliffs, N.J.: Prentice-Hall.

Miller, M. (1980). Executive schemes vs. mental capacity in predicting intellectual underperformance among lower socioeconomic status groups. Unpublished doctoral dissertation, York University, Toronto.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. **Acta Psychologica, 32,** 301-345.

Pascual-Leone, J. (1980). Constructive problems for constructive theories: The current relevance of Piaget's work and a critique of information-processing simulation psychology. In R. Kluwe and H. Spada (Eds.), **Developmental models of thinking.** New York: Academic Press.

Pascual-Leone, J., and Goodman, D. (1979). Intelligence and experience: A neo-Piagetian approach. **Instructional Science, 8,** 301-367.

Pascual-Leone, J., and Smilth, J. (1969). The encoding and decoding of symbols by children: A new experimental paradigm and a neo-Piagetian model. **Journal of Experimental Child Psychology**, 8, 328-355.

Spearman, C. (1927). **The abilities of man.** New York: Macmillan.

Woodcock, R.W. (1980). **Woodcock Language Proficiency Battery, English form.** Hingham, Ma.: Teaching Resources Corp.

Woodcock, R.W. (1981). **Woodcock Language Proficiency Battery, Spanish form.** Hingham, Ma.: Teaching Resources Corp.

Table 1

Mean Independent Variable Scores by Age Group For Full English and Hispanic Samples

	English			Hispanic		
Sex (n)	Male	Female	Tot	Male	Female	Tot
7	9	11	20	8	12	20
9	7	13	20	7	13	20
11	10	10	20	11	9	20

	English Mean	English SD	Hispanic Mean	Hispanic SD
Age				
7	7;9	(4.02)	7;9	(4.82)
9	9;9	(3.95)	9;10	(5.18)
11	11;8	(3.96)	11;8	(4.28)
M-Power				
7	3.1	(0.79)	2.9	(0.97)
9	4.4	(0.77)	4.4	(0.73)
11	5.1	(0.60)	5.0	(0.57)
English Standard-Score				
7	104.9	(10.31)	85.4	(12.77)
9	100.1	(6.99)	83.2	(15.59)
11	101.8	(12.25)	85.4	(12.65)
Spanish Standard-Score				
7			79.2	(13.58)
9			80.9	(14.82)
11			76.9	(15.33)
Teacher Rating				
7			0.64	(1.64)
9			0.96	(1.60)
11			1.36	(1.58)

Table 2

Mean Independent Variable Scores by Age Group for English and Hispanic Subsamples

	English				Hispanic			
	Middle		Working		<= 3 Yrs.		> 3 Yrs.	
Sex (n)	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.
7	4	5	5	6	4	4	4	8
9	3	3	4	10	2	4	5	9
11	7	4	3	6	5	2	6	7
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age								
7	7;8	4.1	7;9	4.1	8;0	4.4	7;6	4.4
9	9;10	2.4	9;9	4.5	9;9	6.0	9;10	5.0
11	11;8	3.3	11;9	4.7	11;8	4.4	11;7	4.3
M-Power								
7	3.1	0.8	3.1	0.8	2.9	1.1	2.9	0.9
9	4.7	0.8	4.3	0.8	4.4	0.6	4.3	0.8
11	5.1	0.7	5.1	0.5	5.1	0.4	5.0	0.7
English Standard-Score								
7	108.0	6.1	102.4	12.5	76.8	10.0	91.2	11.2
9	104.0	6.7	98.4	6.6	71.0	17.6	88.4	11.8
11	108.9	11.3	93.1	6.4	74.0	7.9	91.6	10.2
Spanish Standard-Score								
7					85.1	12.0	75.2	13.6
9					89.8	11.5	77.1	14.8
11					87.9	15.4	71.0	12.1
Teacher Rating								
7					-0.3	1.8	1.3	1.2
9					0.4	1.9	1.2	1.4
11					-0.1	1.6	2.2	0.8

Table 3

Items Used in Metaphor Interpretation Task

English version (1-3, 6, 9 are practice/filler items)

1. The boy ate like a bird.
2. The boy was as quiet as a mouse.
3. The evening breeze is a feather.
4. My sister was a mirror.
5. My shirt was a rock.
6. His smile was a door.
7. My shirt was a butterfly.
8. My sister was a rock.
9. A cloud is a sponge.
10. My shirt was a mirror.
11. My sister was a butterfly.

Spanish version (1-2, 5, 8 are practice/filler items)

1. El hombre trabaja como un burro. (The man works like a burro)
2. El invierno es una nevera. (Winter is a refrigerator)
3. Mi hermana era un espejo. (My sister was a mirror)
4. Mi camisa era una roca. (My shirt was a rock)
5. Su corazon era una jarra. (His heart was a pitcher)
6. Mi camisa era una mariposa. (My shirt was a butterfly)
7. Mi hermana era una roca. (My sister was a rock)
8. El cielo es un mar. (The sky is a sea)
9. Mi camisa era un espejo. (My shirt was a mirror)
10. Mi hermana era una mariposa. (My sister was a butterfly)

Table 4

Across-Ages Correlation Matrices for the English Subsamples

	Metaphor	Age	M-Power	E-Sum'ry	E-Stnd	E-Vocab	E-An/Syn
			Middle Class School n = 26				
Age	78						
M-Power	64	78					
E-Summary	59	87	77				
E-Standard	-17[a]	02[a]	18[a]	51			
E-Vocabulary	50	76	70	90	49		
E-Ant/Syn	58	81	77	95	49	81	
E-Analogy	53	76	58	84	40	59	73
			Working Class Schools n = 34				
Age	74						
M-Power	53	71					
E-Summary	72	72	57				
E-Standard	-06[a]	-39	-19[a]	33			
E-Vocabulary	62	59	47	86	37		
E-Ant/Syn	70	70	56	85	19[a]	60	
E-Analogy	57	62	45	88	29	68	60

Note. All r's significant at $p < .05$, unless noted otherwise.

[a] Not significant.

Table 5

Across-Ages Correlation Matrices for the Hispanic Subsamples

Resident in Canada 3 Years or Less n = 21

	Metaphor	Age	M-Power	E-Sum	E-Stnd	E-Voc	E-An/Sy	E-Anal	S-Sum	S-Stnd	S-Voc	S-An/Sy	S-Anal	E-Teach
Age	86													
M-Power	81	75												
E-Summary	55	63	67											
E-Standard	-20^a	-21^a	05^a	61										
E-Vocabulary	55	66	65	93	48									
E-Ant/Syn	41	47	56	94	71	83								
E-Analogy	54	60	64	89	52	71	77							
S-Summary	64	72	55	38	-25^a	35^a	24^a	43^a						
S-Standard	04^a	03^a	03^a	-07^a	-11^a	-15^a	-08^a	04^a	70					
S-Vocabulary	51	59	33^a	18^a	-35^a	22^a	06^a	20^a	93	72				
S-Ant/Syn	65	74	58	36^a	-29^a	32^a	22^a	44	95	62	82			
S-Analogy	65	70	70	57	-01^a	48	44	64	89	58	69	83		
E-Teacher	-07^a	-04^a	04^a	51	68	41^a	51	53	-12^a	-30^a	-36^a	-17^a	13^a	
S-Story	51	42^a	28^a	04^a	-33^a	-15^a	-05^a	34^a	63	47	57	60	58	

Resident in Canada More Than 3 Years n = 39

	Metaphor	Age	M-Power	E-Sum	E-Stnd	E-Voc	E-An/Sy	E-Anal	S-Sum	S-Stnd	S-Voc	S-An/Sy	S-Anal	E-Teach
Age	78													
M-Power	70	72												
E-Summary	75	78	71											
E-Standard	17^a	-04^a	19^a	58										
E-Vocabulary	69	73	64	92	52									
E-Ant/Syn	71	79	67	94	49	80								
E-Analogy	66	64	65	90	61	74	80							
S-Summary	59	62	56	79	48	73	80	66						
S-Standard	-04^a	-21^a	0^a	22^a	61	20^a	21^a	20^a	63					
S-Vocabulary	45	54	49	72	47	66	74	60	93	63				
S-Ant/Syn	55	55	36	60	26^a	54	62	50	84	50	64			
S-Analogy	67	57	67	78	50	75	75	66	85	50	67	72		
E-Teacher	43	25^a	28^a	55	57	53	48	52	42	28^a	34	34	49	
S-Story	15^a	-01^a	07^a	17^a	26^a	20^a	13^a	13^a	34	44	21^a	45	34	

Note. All r's significant at $p < .05$, unless noted otherwise.

[a] Not significant.

Table 6

Summary Tables for Stepwise Multiple Regression Analyses on English Metaphor Score, by Subsample

Independent Variable	Cumulative R^2	R^2 Change	Cumulative R^2	R^2 Change
		English Sample		
	Working-Class Schools		Middle-Class School	
M-Power	.284	.284	.411	.411
$\overline{E}$-Summary	.543	.259	.453	.042
Age	.616	.073	.638	.185
		Hispanic Sample		
	Resident 3 Yrs. or Less		Resident More than 3 Yrs.	
M-Power	.659	.659	.485	.485
$\overline{E}$-Summary	.659	.000	.613	.127
Age	.814	.155	.676	.063
M-Power	.659	.659	.485	.485
$\overline{S}$-Summary	.710	.051	.543	.058
E-Summary	.710	.000	.613	.070
Age	.814	.104	.676	.063

Table 7

Mean Scores by Age Group for Hispanic Subsample Tested in Spanish

	Age (years)					
	7		9		11	
Sex (n): Male	5		5		5	
Female	5		5		5	
	Mean	SD	Mean	SD	Mean	SD
Age	7;9	(5.07)	9;8	(5.55)	11;9	(4.33)
M-power	2.9	(1.00)	4.4	(0.70)	5.0	(0.67)
English Standard	85.4	(14.08)	84.3	(14.54)	82.90	(13.61)
Spanish Standard	81.2	(17.87)	85.9	(15.04)	80.20	(17.54)
Teacher Rating	0.44	(1.67)	0.80	(1.30)	1.18	(1.89)
Spanish Story	3.5	(0.71)	3.9	(0.57)	3.8	(0.79)
English Metaphor	1.9	(0.87)	3.8	(1.20)	4.9	(0.64)
Spanish Metaphor	2.5	(0.99)	4.0	(0.84)	4.4	(0.79)

Table 8

Across-Ages Correlation Matrices for the Hispanic Subsample Tested in Spanish (*N* = 30)

	Eng. Met.	Spa. Met.	Age	*M*-Power	E-Sum	E-Stnd	E-Voc	E-An/Sy	E-Anal	S-Sum	S-Stnd	S-Voc	S-An/Sy	S-Anal	E-Teach
Span. metaphor	81														
Age	81	66													
M-Power	69	57	69												
E-Summary	67	41	65	57											
E-Standard	05[a]	-15[a]	-16[a]	02[a]	63										
E-Vocabulary	57	36	56	47	93	63									
E-Ant/Syn	61	30[a]	61	51	94	60	83								
E-Analogy	66	47	59	59	83	47	64	70							
S-Summary	52	55	57	57	50	08[a]	35[a]	46	59						
S-Standard	-04[a]	11[a]	-12[a]	13[a]	08[a]	22[a]	-04[a]	05[a]	24[a]	74					
S-Vocabulary	38	46	45	47	39	07[a]	27[a]	37	68	94	76				
S-Ant/Syn	56	55	64	48	48	-02[a]	33[a]	45	55	89	56	72			
S-Analogy	57	53	55	65	55	15[a]	41	48	65	86	60	68	78		
E-Teacher	25[a]	-10[a]	11[a]	14[a]	61	68	61	59	43	04[a]	-07[a]	-06[a]	01[a]	27[a]	
S-Story	12[a]	24[a]	15[a]	26[a]	-05[a]	-19[a]	-12[a]	-07[a]	10[a]	61	61	53	54	61	-10[a]

Note. All *r*'s significant at $p < .05$, unless noted otherwise.

[a] Not significant.

Table 9

Across-Ages Correlations Between Metaphor Scores and Selected Independent Variables for Hispanic Subsample Tested in Spanish. Sample Partitioned by Length of Residence in Canada

	Resident 3 Yrs. or Less n = 13		Resident More than 3 Yrs. n = 17	
	English Metaphor	Spanish Metaphor	English Metaphor	Spanish Metaphor
Spanish Metaphor	71		88	
Age	83	57	80	72
M-Power	74	35[a]	71	67
English-Summary	71	36[a]	72	59
Spanish-Summary	61	71	54	41[a]
Teacher Rating	14[a]	-32[a]	38[a]	18[a]
Spanish Story	42[a]	49[a]	00[a]	03[a]

Note: All r's significant at $p < .05$, unless noted otherwise.

[a] Not significant.

Table 10

Summary Tables for Stepwise Multiple Regression Analyses on Metaphor Scores For Hispanic Subsample Tested in Spanish

	English Metaphor		Spanish Metaphor	
Independent Variable	Cumulative R^2	R^2 Change	Cumulative R^2	R^2 Change
		Full Subsample N = 30		
M-Power	.483	.483	.326	.326
E-Summary	.597	.114	.336	.010
S-Summary	.601	.004	.387	.051
Age	.712	.111	.490	.103
M-Power	.483	.483	.326	.326
S-Summary	.506	.023	.385	.059
E-Summary	.601	.095	.387	.002
Age	.712	.111	.490	.103
		Resident 3 Yrs. or Less n = 13		
M-Power	.545	.545	.122	.122
E-Summary	.599	.054	.143	.021
Age	.744	.145	.333	.190
M-Power	.545	.545	.122	.122
S-Summary	.653	.108	.505	.383
Age	.749	.096	.517	.012
		Resident More than 3 Yrs. n = 17		
M-Power	.498	.498	.449	.449
E-Summary	.601	.103	.477	.028
Age	.701	.100	.568	.091
M-Power	.498	.498	.449	.449
S-Summary	.512	.014	.449	.000
Age	.682	.170	.575	.126

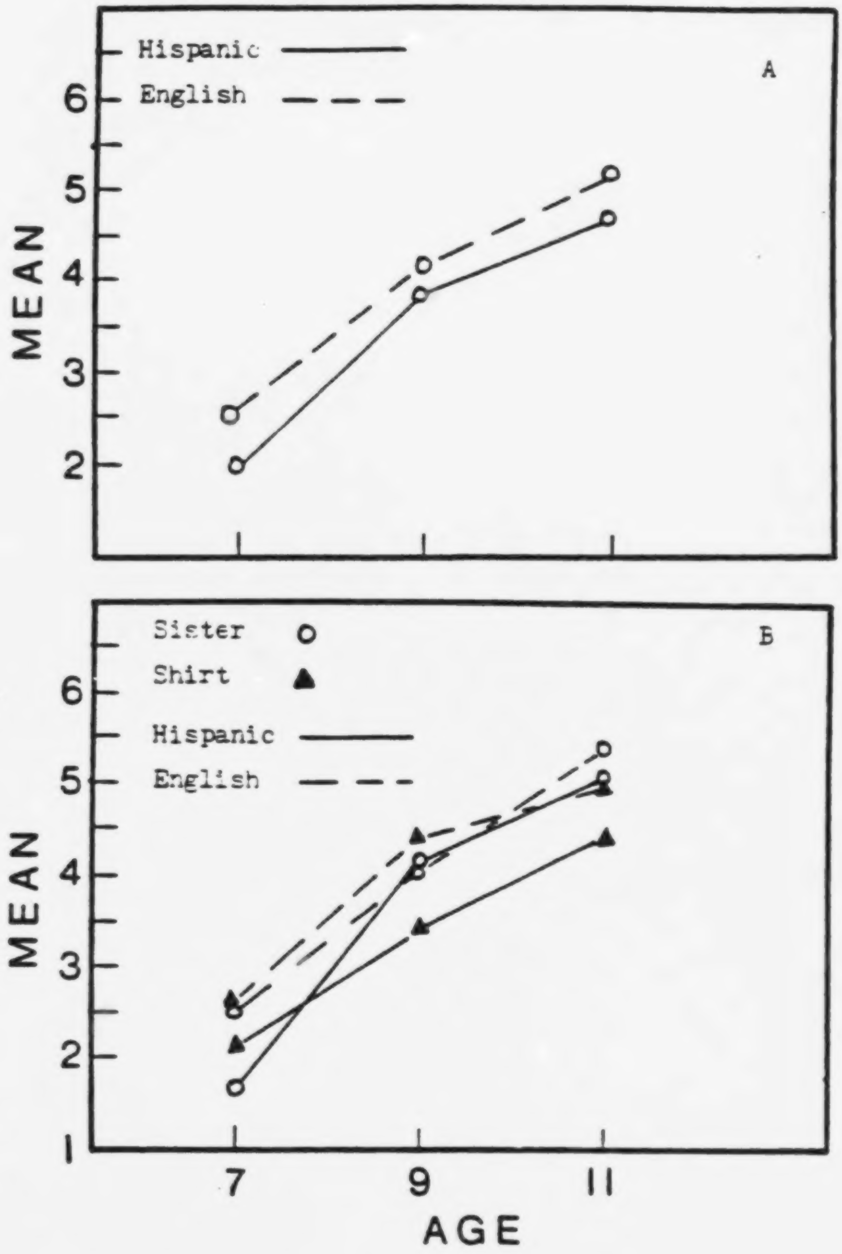


Figure 1

Full Hispanic and English samples. A: Mean metaphor point score as a function of age and language group. B: Mean metaphor point score as a function of age, language group, and topic

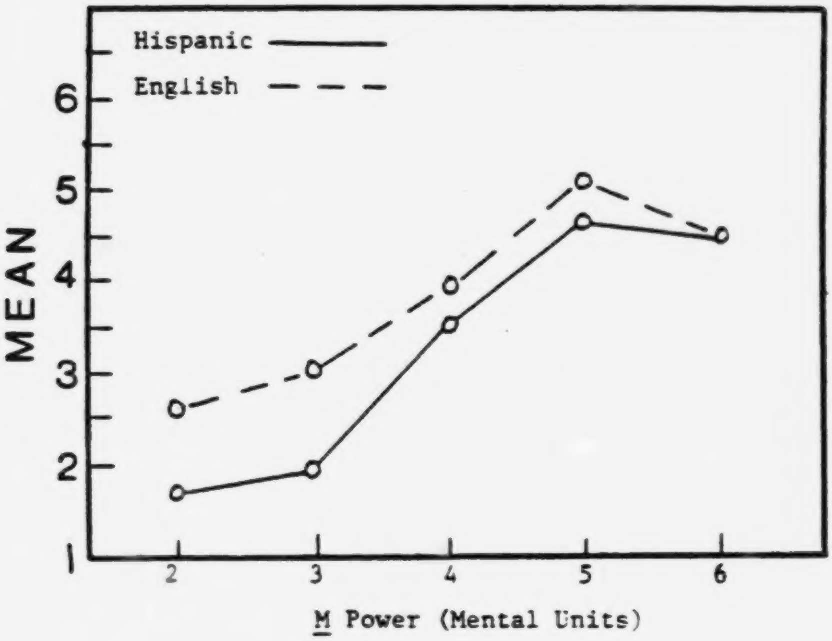


Figure 2

Full Hispanic and English samples. Mean metaphor point score as a function of measured M-power and language group

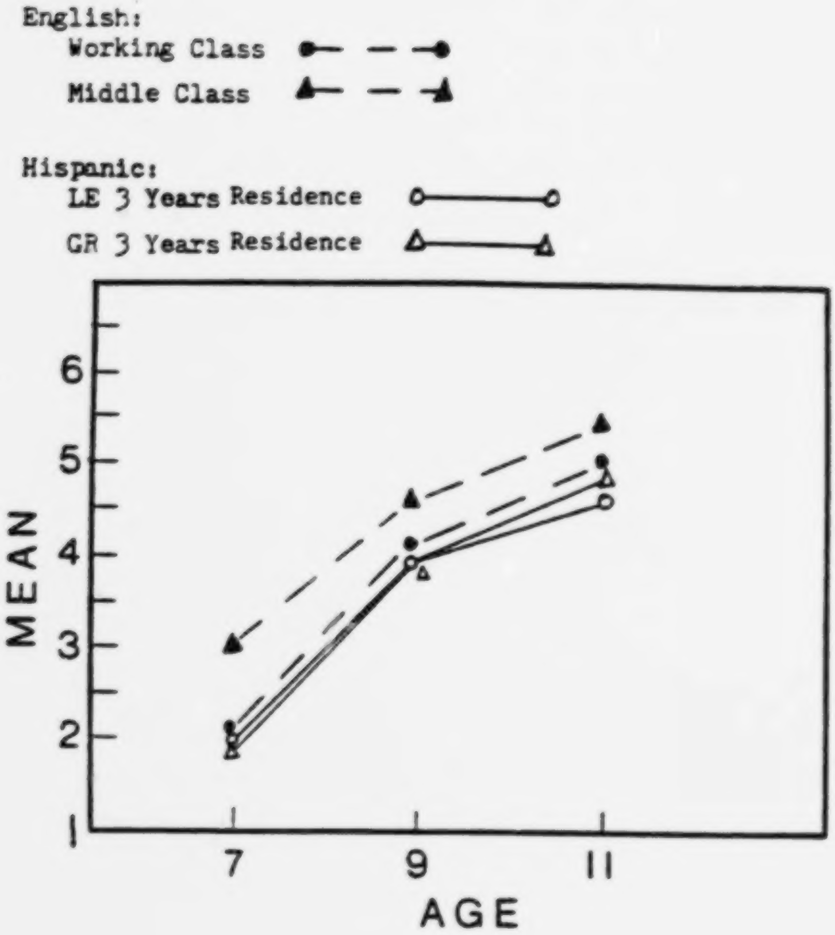


Figure 3

Hispanic and English subsamples. Mean metaphor point score as a function of age and subsample

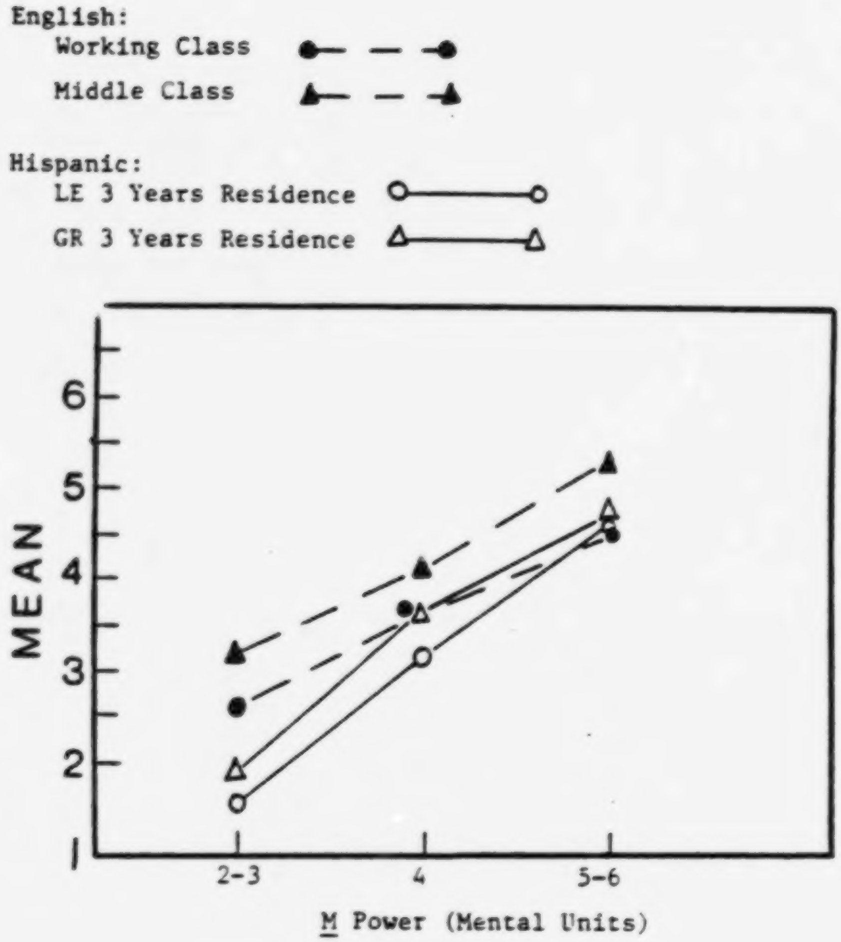


Figure 4

Hispanic and English subsamples. Mean metaphor point score as a function of measured M-power, and subsample

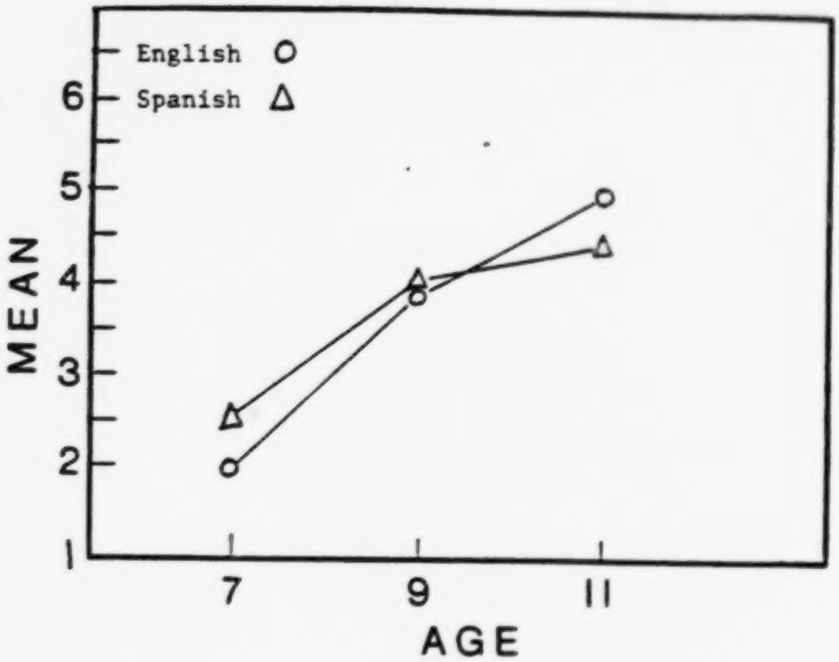


Figure 5

Hispanic subsample tested in English and Spanish. Mean metaphor point score as a function of age and test language

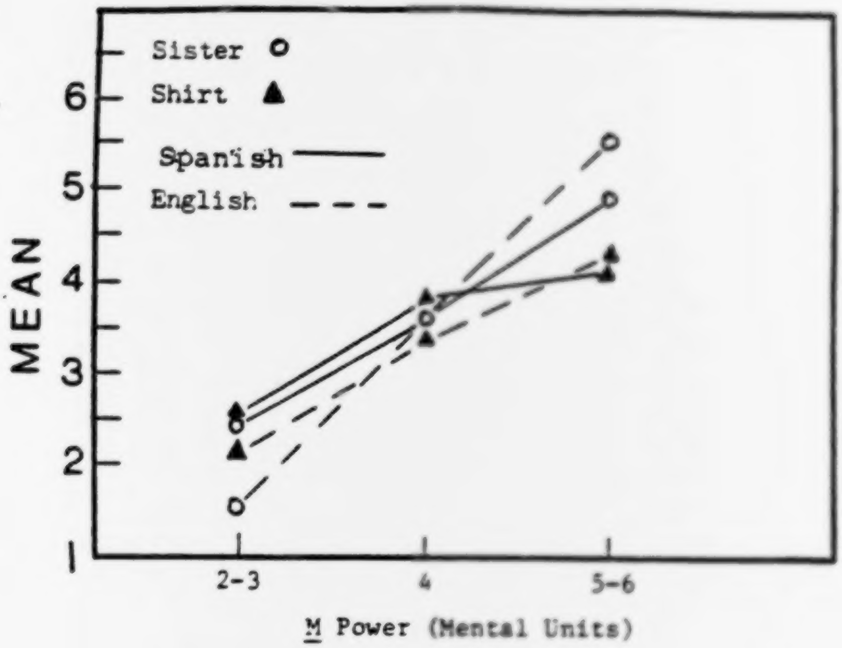


Figure 6

Hispanic subsample tested in English and Spanish. Mean metaphor point score as a function of measured M-power, test language, and topic

Appendix

1. Language questionnaire completed by English parents
2. Language questionnaire completed by Hispanic parents
3. Table A1. Sample items from Oral Language subscale of the **Woodcock** (1980) **Language Proficiency Battery.**
4. Rating form completed by teachers for Hispanic subjects
5. Table A2. Across-ages correlation matrices for full English and Hispanic samples.

Please sign and return to the School by February 25

☐ I give my permission for ____________________ to participate
(student's name)

☐ I do not give my permission in the "Metaphor" research project.

______________________________ ______________
(signature of parent or guardian) (date)

Since we are studying children from various language groups, we need information on the language background of the children. If you are giving permission for your child to participate, please complete the following questionnnaire. For questions 2 to 5, please write-in the language primarily spoken (first part of each question) and any other languages spoken (second part). For each language, if it is the only language spoken in the situation indicated by the question, check the box marked "always"; if the language is spoken hardly at all, check the box marked "rarely"; if the use is somewhere between these extremes, check the appropriate intermediate box.

1. What language was first learned by your child? ____________________

	HOW MUCH OF THE TIME				
	rarely	some-times	half the time	often	always
2. What language is primarily spoken by your child at home? ____________	☐	☐	☐	☐	☐
Any other language/s? ____________	☐	☐	☐	☐	☐
3. What language is primarily spoken by adults in the child's home? ____________	☐	☐	☐	☐	☐
Any other language/s? ____________	☐	☐	☐	☐	☐
4. What language is primarily spoken by other children in the child's home? ____________	☐	☐	☐	☐	☐
Any other language/s? ____________	☐	☐	☐	☐	☐
5. What language is primarily spoken by your child in informal social situations (playground, cafeteria, or on the street)? ____________	☐	☐	☐	☐	☐
Any other language/s? ____________	☐	☐	☐	☐	☐

THANK YOU

Firme, por favor, y devuelva a la escuela en o antes del 25 febrero

☐ Doy mi permiso para que ______________________ participe
(nombre del niño/niña)

☐ No doy mi permiso en la investigación sobre metáforas.

______________________ ____________
(firma del padre o madre o tutor) (fecha)

Si Vd. permite la participación de su hijo/hija llene, por favor, el questionario siguiente acerca de las lenguas (idiomas) que su hijo/hija utiliza. En las preguntas de números 2 a 5, escriba primero acerca de la lengua utilizada mas frecuentemente y luego (segunda línea) de otras lenguas. Para cada lengua marque la casilla apropiada. Si es el única lengua usada en la situación que la pregunta describe, marque la casilla de "continuamente"; si la lengua no se usa casi nada, marque la casilla de "raramente"; si el uso es entre estos extremos, marque la casilla apropiada.

1. ¿Cual fue la lengua materna (primera) de su hijo/hija? ______________________

	CUANTO TIEMPO				
	Raramente	A veces	A tiempo medio	A menudo	Continuamente
2. ¿Que lengua (idioma) habla el hijo/hija corrientemente en su hogar? ______________	☐	☐	☐	☐	☐
¿Otra(s) lengua(s)? ______________	☐	☐	☐	☐	☐
3. ¿Que lengua hablan los adultos corrientemente en su hogar? ______________	☐	☐	☐	☐	☐
¿Otra(s) lengua(s)? ______________	☐	☐	☐	☐	☐
4. ¿Que lengua hablan otros niños corrientemente en el hogar se su hijo/hija? ______________	☐	☐	☐	☐	☐
¿Otra(s) lengua(s)? ______________	☐	☐	☐	☐	☐
5. ¿Que lengua habla su hijo/hija corrientemente en situaciones de la vida diaria (en la calle, en la cafetería, en los juegos)? ______________	☐	☐	☐	☐	☐
¿Otra(s) lengua(s)? ______________	☐	☐	☐	☐	☐

MUCHAS GRACIAS

Table A1

Sample Items from the Oral Language Subscale of the Woodcock (1980) Language Proficiency Battery

Picture Vocabulary

The subject is asked to provide names for pictured objects or actions. The child is shown a sequence of pictures and asked to verbally label each picture. One easy and one more difficult item:

No.2 (Subject is shown a picture of scissors), E points to picture and says "What is this?"
Correct: scissors, shears

No. 26 (Subject is shown a picture of hinges), E points to picture and says "What are these?"
Correct: hinges. **Incorrect:** clamps, latches.
Query: door holders, something for doors — "What are they called?"

Antonyms/Synonyms

In Part A, Antonyms, the subject must state a word whose meaning is the opposite of a given word. In Part B, Synonyms, the subject must state a word whose meaning is approximately the same as a given word.

Antonyms:
No.1. Point to the first word on the subject's side and say "Tell me the opposite of 'down'." **Correct:** up.

No. 19 Point to the first word on the subject's side and say "Tell me the opposite of 'authentic'."
Correct: fake, counterfeit, false, fraudulent, imitation, phony, untrue.
Incorrect: artificial, duplicate, synthetic.
Query: copy, forgery, replica, reproduction — "Tell me another word."

Synonyms:
No.2. "Tell me another word for 'lawn'."
Correct: grass, sod, yard. **Incorrect:** cut, mow.

No.20. Point to the first word on the subject's side and say "Tell me another word for 'chide'."
Correct: scold, rebuke, reprimand, reprove.
Incorrect: kid, ridicule, tease.
Query: correct, punish — "Tell me another word."

Table A1, continued

Analogies

The subject must complete oral statements of verbal analogies.

No.1 Point to the first item on the subject's side and say: "Scissors is to cut as pencil is to ... " (pause).
Correct: write, color, draw, mark, record.
Query: pen — "Tell me another word."

No.23. "Victory is to winner as diploma is to ... " (pause).
Correct: graduate
Incorrect: college, graduation, loser, scholar, student.

LAS® OBSERVATION FORM

ENGLISH

Student's name: ____________________

Person filling out form: ____________________

Date: ____________________

Based on your personal knowledge, to what extent would this student have difficulty in the following situations? Note that in filling out the seven point scale, 0 (the midpoint) should only be used for a neutral or "don't know" opinion.

	Difficulty — No Difficulty
1. Explaining in English to a monolingual playmate how to play a game.	-3 -2 -1 0 +1 +2 +3
2. Giving directions in English for getting to a friend's house from school.	-3 -2 -1 0 +1 +2 +3
3. Explaining in English to a teacher why he or she did not complete a homework assignment.	-3 -2 -1 0 +1 +2 +3
4. Giving an oral report in English to the class.	-3 -2 -1 0 +1 +2 +3
5. Delivering an oral message in English from the teacher to a monolingual English speaking secretary.	-3 -2 -1 0 +1 +2 +3

Based on your personal knowledge rate the child in the following oral language areas.

	Low Skill — High Skill
6. English pronunciation:	-3 -2 -1 0 +1 +2 +3
7. English comprehension:	-3 -2 -1 0 +1 +2 +3
8. English vocabulary:	-3 -2 -1 0 +1 +2 +3
9. English syntax:	-3 -2 -1 0 +1 +2 +3
10. General communicative competence in English:	-3 -2 -1 0 +1 +2 +3

How long have you known this student? ____________

This form may be reproduced as necessary.

Table A2

Across-Ages Correlation Matrices for Full English and Hispanic Samples

English Sample n = 60

	Metaphor	Age	M-Power	E-Sum'ry	E-Stnd	E-Vocab	E-An/Syn
Age	75						
M-Power	58	74					
E-Summary	67	76	62				
E-Standard	02[a]	-16[a]	-03[a]	51			
E-Vocabulary	58	67	55	88	46		
E-Ant/Syn	66	72	62	92	44	73	
E-Analogy	58	66	49	88	45	65	72

Hispanic Sample n = 60

	Metaphor	Age	M-Power	E-Sum	E-Stnd	E-Voc	E-An/Sy	E-Anal	S-Sum	S-Stnd	S-Voc	S-An/Sy	S-Anal
Age	81												
M-Power	73	73											
E-Summary	64	66	62										
E-Standard	08[a]	08[a]	11[a]	68									
E-Vocabulary	59	61	55	94	65								
E-Ant/Syn	59	63	57	95	66	84							
E-Analogy	62	61	63	88	59	72	79						
S-Summary	55	62	52	41	-03[a]	31	38	46					
S-Standard	04[a]	-12[a]	01[a]	-08[a]	02[a]	-15[a]	-08[a]	04[a]	70				
S-Vocabulary	38	49	39	26	-11[a]	17[a]	25	31	93	72			
S-Ant/Syn	55	61	44	36	-08[a]	28	33	41	88	57	70		
S-Analogy	64	61	67	59	19[a]	51	53	61	83	49	62	76	
E-Teacher	23[a]	12[a]	16[a]	63	72	61	57	55	-03[a]	-15[a]	-16[a]	-03[a]	24[a]

Note. All r's significant at p <.05, unless noted otherwise.

[a] Not significant.

THE DEVELOPMENT OF BILINGUAL PROFICIENCY

FINAL REPORT

Volume II: Classroom Treatment

Birgit Harley
Patrick Allen
Jim Cummins
Merrill Swain

April 1987

Modern Language Centre
The Ontario Institute for Studies in Education

TABLE OF CONTENTS

Page

VOLUME I
Preface vii
Project Staff viii

Introduction
THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. Purpose 1
2. The Nature of Language Proficiency 1
3. Classroom Treatment Studies 8
4. Social Context and Age 13
5. Conclusions 19

Chapter 1
PERSPECTIVES ON LEXICAL PROFICIENCY IN A SECOND LANGUAGE

1. Introduction 23
2. The Data Base 24
3. Measures of Lexical Proficiency 26
4. Lexical Measures and L2 Proficiency 30
5. Lexical Use of Immersion Students and Native Speakers 34
6. Conclusion 40

Tables 44
Figures 60
Appendix A 64
Appendix B 66

Chapter 2
THE DEVELOPMENTAL GROWTH OF METAPHOR COMPREHENSION IN CHILDREN'S FIRST AND SECOND LANGUAGE

1. The Development of Bilingualism 67
2. A Cognitive Developmental Perspective 69
3. Summary of the Present Study and Hypotheses 70
4. Method 71
5. Results: Metaphors Interpreted in English 76
6. Results: Metaphor Interpretations in Spanish 81
7. Conclusions 84

Tables 91
Figures 101
Appendix 107

Page

VOLUME II
Preface vii
Project Staff viii

Introduction
THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. Purpose 1
2. The Nature of Language Proficiency 1
3. Classroom Treatment Studies 8
4. Social Context and Age 13
5. Conclusions 19

Chapter 3
COLT OBSERVATION SCHEME: DEVELOPMENT AND VALIDATION

1. Introduction 23
2. Description and Rationale of the Observation Scheme 24
3. The Validation Study 36
4. Discussion 43

Tables 50
Figures 52
Appendix A 54

Chapter 4
THE CORE FRENCH OBSERVATION STUDY

1. Introduction 56
2. The Sample 57
3. General Procedures 59
4. The Colt Observation Scheme: Procedures 61
5. Classroom Practice: Findings 64
6. Proficiency Predictions 70
7. Proficiency Tests 72
8. Process/Product Findings 78
9. Discussion and Interpretation of Results 86

Tables 98
Appendix A 127
Appendix B 137
Appendix C 161
Appendix D 164
Appendix E 167
Appendix F 181

Page

VOLUME II (continued)

Chapter 5
THE IMMERSION OBSERVATION STUDY

1. Introduction 190
2. Vocabulary Instruction in Immersion Classes 192
3. Vous/Tu Input 222
4. Student Talk in Teacher-Fronted Activities 226
5. Error Treatment 233
6. Concluding Comment 240

Tables 252
Appendix A 264
Appendix B 287
Appendix C 293
Appendix D 317

Chapter 6
FUNCTIONAL GRAMMAR IN FRENCH IMMERSION: A CLASSROOM EXPERIMENT

1. Purpose and Rationale 342
2. Design of the Experiment 346
3. The Findings 354
4. Discussion 361
5. Conclusion 363

Tables 367
Figures 379
Appendix A 381
Appendix B 385
Appendix C 386
Appendix D 389
Appendix E 390
Appendix F 391
Appendix G 399
Appendix H 409

Page

VOLUME III

Preface vii
Project Staff viii

Introduction
THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. Purpose 1
2. The Nature of Language Proficiency 1
3. Classroom Treatment Studies 8
4. Social Context and Age 13
5. Conclusions 19

Chapter 7
THE LANGUAGE USE PATTERNS, LANGUAGE ATTITUDES, AND BILINGUAL PROFICIENCY OF PORTUGUESE CANADIAN CHILDREN IN TORONTO

1. Introduction 23
2. Scoring Procedures and Refinement of the Data Set 28
3. Results 35
4. Comparison of Portuguese Proficiency with Azorean Native Speakers 49

Tables 54
Appendix A 85
Appendix B 88
Appendix C 106
Appendix D 119
Appendix E 154

Chapter 8
THE DEVELOPMENT OF BILINGUAL PROFICIENCY IN THE TRANSITION FROM HOME TO SCHOOL: A LONGITUDINAL STUDY OF PORTUGUESE-SPEAKING CHILDREN

1. Purpose of the Study 159
2. Methodology 159
3. Scoring Procedures and Preliminary Results for Home Interaction Data 165
4. Home Interaction of 4-Year Old Portuguese Background Children in Toronto: Preliminary Trends 176

Tables 182

Page

VOLUME III (continued)

Chapter 9
AGE OF ARRIVAL, LENGTH OF RESIDENCE, AND INTERDEPENDENCE OF LITERACY SKILLS AMONG JAPANESE IMMIGRANT STUDENTS

1. Background 183
2. Methodology 185
3. Results 188
4. Conclusion 191

Tables 194
Figures 199
Appendix A 200

Chapter 10
THE RELATIONSHIP BETWEEN STARTING AGE AND ORAL SECOND LANGUAGE PROFICIENCY IN THREE GROUPS OF CLASSROOM LEARNERS

1. Introduction 203
2. The Sample 204
3. Procedures 205
4. Verb Use: Analysis and Results 207
5. Fluency: Analysis and Results 213
6. Sociolinguistic Test Results 216
7. Discussion 216
8. Conclusion 220

Tables 223
Appendix A 232
Appendix B 235
Appendix C 253
Appendix D 251

PREFACE

The Development of Bilingual Proficiency is a large-scale, five-year research project which began in September 1981. The present Final Report of the project is the third in a series. It follows an interim Year 1 Report, produced in September 1982 at the end of the first year of the project, and a Year 2 Report, produced in September 1983.

There are three volumes in this Final Report of the project, each concentrating on specific issues investigated in the research: the nature of language proficiency (Volume I), the effect of classroom treatment on language proficiency (Volume II), and the relevance of social context and age for language learning (Volume III). Each volume is introduced by an identical 20-page overview of all the studies carried out in the context of the Development of Bilingual Proficiency (DBP) Project. The overview includes brief summaries of the individual studies together with an indication as to where the complete report of each study is to be found (either in the Year 2 Report or in Volume I, II, or III of the Final Report). Within the complete reports of individual studies contained in this Final Report, references to other Development of Bilingual Proficiency Project studies appear either as 'Year 2 Report' or, when they form part of the Final Report, as chapter numbers only. Note that Chapters 1 and 2 appear in Volume I, Chapters 3-6 in Volume II, and Chapters 7-10 in Volume III.

The authors wish to acknowledge the contribution of the many individuals and organizations who have played a role in the Development of Bilingual Proficiency Project since it began in September 1981. We are grateful to administrators, teachers, students and their parents in the following Ontario school boards for their participation in the research: the Board of Education for the City of Scarborough, the Carleton Board of Education, the Metropolitan Separate School Board, the North York Board of Education, and the Toronto Board of Education. In addition, we would like to thank the Portuguese Secretary of State for Immigration, the Regional Secretary of Social Affairs for the Autonomous Region of the Azores, and the staff, parents and students of the Japanese School of Toronto Shokokai Inc. We wish to acknowledge the contribution of Ellen Bialystok and Raymond Mougeon, who were principal investigators of the project in 1981-82 and 1981-83 respectively. We would also like to express our appreciation to project staff for their part in carrying out the research and in text-processing. Finally, we wish to acknowledge the financial support provided in the form of a five-year negotiated grant (No. 431-79-0003) by the Social Sciences and Humanities Research Council of Canada, and the administrative and financial contribution of the Ontario Institute for Studies in Education.

B.H., P.A., J.C., M.S.

PROJECT STAFF

(from September 1983)*

Principal Investigators:

Patrick Allen, Jim Cummins, Birgit Harley, Merrill Swain

Project Managers:

Birgit Harley, Sharon Lapkin

Research Officers:

Margarida Aguiar, Jud Burtis, Susanne Carroll, Maria Fröhlich, Vince Gaudino, Mary Lou King, Laurette Lévy, José Lopes, Françoise Pelletier, Bruno Piché, Jorge Ramos, Mireille Tremblay, Virginia Won (Pang)

Graduate Assistants:

Jane Appelt, Eva Bild, Patrick Conteh, Giselle Corbeil, Vera Held, Richard Kruk, Bill Marshall, Josée Mazzilli, Anne Riddle, Emilia Rivas-Arraiz, Kathryn Shred, Nina Spada, Yilin Sun, Tim Tomlinson

Post-doctoral Fellows:

Monica Heller, Janice Johnson

Statistical Consultant:

Jamshid Etezadi-Amoli

Text-processing Specialist:

Ellen Jeske

Secretaries:

Joyce Howlett, Doreen Panowyk, Donna Schillaci

* For staff members in previous years, see earlier reports

Introduction

THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. PURPOSE

The purpose of this five-year research project has been to investigate issues concerning language proficiency[1] and its development in educational contexts for children learning a second language. The research has concentrated on the following major issues: the nature of language proficiency; the impact of instructional practices on language learning; the relationship between social-environmental factors and bilingual proficiency; and the relationship between age and language proficiency. In this overview of the project, studies focussing on each of these issues are summarized.

2. THE NATURE OF LANGUAGE PROFICIENCY

The focus and ultimate goal of all studies carried out within the Development of Bilingual Proficiency Project is the improvement of educational practices as they relate to second language learning and teaching. Because so much of school practice relates rather narrowly to the teaching and learning of grammatical aspects of the target language, it was considered essential to broaden the scope of the typical educational definition of language proficiency to incorporate discourse and sociolinguistic dimensions, and to consider the differential demands that context-reduced versus more context-embedded language tasks may make on the learner.

2:1 Large-scale Proficiency Study
(Year 2 Report)

Objectives. The primary purpose of the large-scale proficiency study conducted during the first two years of the Project was to determine whether the three hypothesized traits, representing key components of language proficiency, could be empirically distinguished. It was hypothesized that grammatical, discourse, and sociolinguistic competence would emerge as distinct components of second language proficiency which may be differentially manifested under different task conditions. A secondary purpose of the study was to develop a set of exemplary test items and scoring procedures that could be used, or modified for use, in further studies involving the measurement of the hypothesized traits. A final purpose of the study was to provide a broadly based description of the target language proficiency of the second language learners tested, in relation to that of native speakers.

Subjects. A total of 198 students was involved in the study. Of these, 175 were grade 6 early French immersion students from the Ottawa region, and 23 were grade 6 native speakers from a regular Francophone school in Montreal. The immersion students, in six intact classes, had received 100% of their schooling in French in kindergarten to grade 2 or 3, since when they had been taught in English for a gradually increasing portion of each day. At the time of testing, about 50% of their school subjects were

being taught in French, and the other 50% in English. This sample of classroom second language learners was selected because of the theoretically interesting and educationally innovative nature of their intensive school-based language learning experience, and because they were at an age where they were sufficiently proficient in the second language to be able to cope with a wide range of types of language tasks.

Instruments. A multi-method multi-trait design was used to determine the extent to which grammatical, discourse, and sociolinguistic dimensions of the immersion students' French proficiency were distinguishable. To measure proficiency on each trait, three methods of testing were used: oral production, multiple choice, and written production. A matrix with nine test cells was thus created, consisting of three tests of grammar, three of discourse, and three of sociolinguistics. The oral production task for each trait was administered to a randomly selected sub-sample of 69 immersion students and ten native speakers, representing ten-eleven subjects from each class.

Grammatical competence was operationalized for the purposes of this study as rules of morphology and syntax, with a major emphasis on verbs and prepositions. The grammar oral production task consisted of a guided individual interview in which the interviewers' questions were designed to elicit a variety of verb forms and prepositions in French, as well as responses that were sufficiently elaborated to score for syntactic accuracy. The content of the interview questions (e.g. favourite pastimes, trips taken) was at the same time designed to focus the subject's attention on communication rather than the code. Grammatical scoring was based on the student's ability to use certain grammatical forms accurately in the context of particular questions. The group-administered grammar multiple choice test consisted of 45 written items which also assessed knowledge of the verb system, prepositions, and other syntactic rules, including written agreement rules. The student's task was to select the correct response from three alternatives provided. The third grammar task, written production, consisted of two short compositions to be written in 15 minutes each — one a narrative and the other a letter of request. Both this written production task and a parallel discourse written production task — also involving a narrative and a request letter — were assessed for grammatical proficiency. Scoring focussed on grammatical accuracy in verbs, prepositions, and other rules of syntax and morphology.

The discourse trait was defined as the ability to produce and recognize coherent and cohesive text (written or oral). For the individual discourse oral production task, the student was required to retell the story of a silent movie and to present arguments in support of an opinion. This task was rated on 5-point scales both globally and in detail for coherence and cohesion, focussing, for example, on the student's ability to make clear and accurate reference to characters, objects, and locations, to produce a logically connected text, and to fulfill the basic task required. The discourse multiple choice task consisted of 29 short written passages from each of which a sentence had been omitted. The student was required to select from three alternatives the sentence that best fit the context. The discourse written production task, like the grammar written production task, consisted of a narrative and a request letter. All four (grammar and discourse) tasks were rated for proficiency in discourse on the same kinds of features that were assessed in the discourse oral production task.

Sociolinguistic competence, the third trait dealt with in this proficiency study, was operationalized as the ability to produce and recognize socially appropriate language in context. The individual oral production task involved a set of slides with taped descriptions representing situations of different levels of formality. The student's task

was to respond appropriately with a request, offer, or complaint. Scoring focussed on the student's ability to shift register appropriately. Thus sociolinguistic proficiency was measured by difference scores, calculated by subtracting the number of formal 'politeness' markers produced by the student in informal variants of situations from those produced in formal variants of the situations. The sociolinguistic multiple choice test consisted of 28 items, each with three alternative ways of expressing a given sociocultural function. The choices were all grammatically accurate but not equally appropriate. The student's task was to select the most appropriate of the choices in the given situation. Scoring of responses was weighted according to the choices made by native speakers. The sociolinguistic written production task involved the writing of a formal request letter and two informal notes, all of which could be categorized as directives. The request letter written as part of the discourse written production task was also scored for sociolinguistic proficiency. As for the oral production task, scoring was based on difference scores, calculated by subtracting the number of formal markers produced in the notes from those produced in the letters.

Reliability and generalizability of scores. The component within-test scores were combined to produce a single overall score for each of the nine trait-method cells in the matrix. The composition of each of these overall scores was calculated to maximize validity and reliability. On the multiple choice tests, the reliability of the immersion students' total scores ranged from .58 on the sociolinguistic test to .75 on the discourse test. Generalizability studies were conducted on those cells for which sufficient data were available: the sociolinguistic oral production test and the three written production tests. G-coefficients for these tests, based on the subsample of orally tested students, were comparable to the multiple choice test reliabilities.

Testing a model of proficiency. In order to determine whether the three traits — grammatical, discourse, and sociolinguistic competence — could be empirically distinguished, two kinds of analyses were performed: (a) factor analysis, and (b) a comparison of the group means of the learners and native speakers.

The factor analysis based on the 69 orally tested immersion students failed to confirm the hypothesized three-trait structure of proficiency. Instead, confirmatory factor analysis by means of LISREL produced a two-factor solution. One of these factors was interpretable as a general language proficiency factor; it had positive loadings from all cells in the nine-test matrix except for the sociolinguistic written production test. The highest loadings on this general factor were from the three grammatical tests. The second factor was interpretable as a written method factor; it had loadings from the three multiple choice tests and from all three written production tests. The tests loading on this method factor appeared to be tapping the kind of literacy-oriented linguistic proficiency that is typically learned in classrooms. The lack of trait structure emerging from the factor analysis may have been due to the homogeneity of the immersion student sample. Within their classroom setting these students had all had very much the same kind of exposure to French, and strong opportunities for some students to develop proficiency in one area, and other students to develop proficiency in a different area, were lacking.

A different kind of result emerged from comparisons of immersion and native-speaker scores on the various tests. On all three grammar tests, the immersion students' mean score was considerably lower than that of the native speakers (p .01), and they also scored generally lower on the sociolinguistic tests than did the native speakers. On the discourse tasks, however, the scores of the immersion students were close or

equivalent to those of the native speakers and there were no significant between-group differences. The immersion students' strong performance in discourse may have been due to positive transfer from prior experience in their mother tongue. In contrast to the factor analysis results, then, these comparative findings, showing very different results for discourse as opposed to grammar and sociolinguistics, provide some evidence in support of a distinction between traits.

Conclusions. It was concluded that, although the three hypothesized language proficiency traits were not empirically distinguished via the factor analysis, this result may have been dependent on the relatively homogeneous language learning background of the immersion population studied. This did not necessarily mean that the traits would not be distinguishable in a more heterogeneous language learning population. From an educational perspective it was clear that the analysis of proficiency into different components was diagnostically revealing of the second language strengths and weaknesses of the immersion students. It was decided that two kinds of further studies were indicated to probe issues concerning how different dimensions of proficiency develop as a function of the immersion students' specific language learning experience: (a) small-scale in-depth studies of specific aspects of the immersion students' second language proficiency based on the data already collected (see 2:2 - 2:3 below), and (b) the study of language learning activities in the immersion classroom setting (see 3:3 - 3:4 below).

2:2 Transfer in Immersion Students' Compositions (Year 2 Report)

Hypotheses and design. Given the shared mother tongue, English, of the immersion students and the dominance of English in the wider school and outside-school environment of the immersion program, mother tongue transfer was expected to be a continuing factor in the students' written production at the grade 6 level. In a small-scale study of compositions written by 22 native speakers and 22 of the orally tested immersion students from two randomly selected classes in the larger proficiency study, evidence was sought for the hypothesis that mother tongue transfer may be manifested in the way in which the learners were distributing semantic information across syntactic elements in the second language, without necessarily making outright errors.

One of the composition topics assigned in the large-scale proficiency study, Au secours!, involved writing a story about the rescue of a kitten from a tree. The students' stories on this topic contained a very similar series of events, involving several changes of location. The focus of the present study was on how the immersion students were expressing the location/direction distinction in these stories, given that there are characteristic differences between French and English in this linguistic domain. While in English, prepositions generally serve an important role in conveying the location/direction distinction (e.g. at/to, in/into), in French there is a general tendency for direction to be expressed in the verb, and for prepositions (e.g. à, dans, sur) to be neutral with respect to the location/direction distinction. It was hypothesized that the immersion students would rely on prepositions rather than the verb to express the notion of direction.

Findings. A comparison of directional expressions in the Au secours! stories written by the immersion students and the native speakers showed that, as expected, the immersion students were much less likely than the native speakers to mark direction in the verb, preferring a non-directional verb of motion such as courir together with a

preposition to express the directional notion. The immersion students, at the same time, sometimes erroneously used French prepositions unmarked for direction as if they were carrying the directional distinction, and also tended to make more use than the native speakers of prepositional phrases expressing direction, even on those occasions when they also used directional verbs. This latter tendency did not necessarily lead to error. The findings of the study thus provide support for the hypothesis that the immersion students would show a systematic tendency to rely more heavily on prepositions to express the notion of direction than the native speakers.

Conclusions. It was concluded that the students may need more focussed classroom input that would alert them to such characteristics of French that are different from English, together with more opportunities for expressing the relevant distinctions in their second language.

2:3 Lexical Proficiency in a Second Language
(Final Report, Vol. I)

In the large-scale proficiency study described above (2:1), there were no measures specifically designed to assess lexical proficiency, not because lexical proficiency was considered unimportant but because it was assumed to enter into performance on all the tasks assigned. In the present study, the two narratives and three request letters written by 69 immersion students and 22 native speakers in the context of the various written production tests were re-analysed from a lexical perspective, with verbs being selected as the focus for the study. The purpose of the study was threefold: (a) to compare different quantitative measures of immersion students' lexical proficiency in their second language (L2) writing; (b) to examine the relationship between written lexical proficiency and other aspects of their L2 communicative competence, and (c) to describe the students' lexical use in relation to that of native speakers.

(a) Measures of lexical proficiency. Five quantitative measures of lexical proficiency were developed and statistically compared. One of these was a 'lexical error rate', while the other four were variations on the theme of lexical richness, labelled respectively 'number of lexical types', 'lexical variety', 'lexical specificity', and 'lexical sophistication'. All the measures, except for 'number of lexical types' were controlled for length of text. For each student the data from the five written compositions were lumped together. Two of the relatively difficult measures were retained as the most appropriate for further use in a factor analysis. The first was 'lexical specificity', which consisted of the number of different verb types used by each student, not counting the 20 most frequent verbs in French or those that were used in the instructions to the compositions, divided by the number of verb items produced. The second measure was 'lexical sophistication', representing those relatively infrequent verbs not found in a basic word frequency list, also divided by the number of verb items produced.

(b) Lexical measures and L2 proficiency. Three mutually exclusive hypotheses arising from previous work were examined via factor analysis: (1) that lexical proficiency is equally involved in all three of the components of language proficiency examined in the large-scale proficiency study: namely, grammar, discourse, and sociolinguistics; (2) that lexical proficiency is part of the grammar component; or (3) that lexical proficiency is a separate component, distinct from the other three components of language proficiency.

Confirmatory factor analyses showed that an acceptable solution to fit any one of these three hypotheses could be found, and that there was no conclusive evidence favouring any one of the three hypotheses over the other two. One interesting finding was that in the three- and four-factor solutions corresponding to hypotheses (2) and (3) respectively, a grammar factor and a discourse factor emerged, which had not previously been found in the large-scale study where no lexical measures had been included.

(c) Lexical use of immersion students and native speakers. A comparison of the verb lexis used by the immersion students and the native speakers in their compositions revealed that the immersion students tended to make proportionately more use of high-coverage verbs, and less use of some morphologically or syntactically complex verbs such as pronominal and derived verbs. The inflectional complexity of some high coverage verbs did not appear to be a deterrent to their use although inflectional errors (considered grammatical rather than lexical errors) did occur. Semantic and syntactic incongruence with their English mother tongue (L1) emerged as an important factor in the immersion students' non-use of some French verb types and in the lexical errors they made. At the same time, the students demonstrated positive L1 transfer in the use of some cognate verbs in French.

Conclusions. It was suggested that the immersion students' stock of lexical items would benefit from more classroom activities designed to increase their use of L2 derivational resources and to emphasize the use of more specific vocabulary.

2:4 Communicative Skills of Young L2 Learners (Year 2 Report)

Purpose and data base. This exploratory study involved a detailed investigation of methods of scoring oral L2 performance and of the interrelationships among various aspects of L2 proficiency. The study was based on a subset of data previously collected in the context of another Modern Language Centre project. It consisted of oral tasks in English with 22 Japanese immigrant students in grades 2, 3, 5 and 6, together with academic tests of reading and vocabulary in the L2.

Findings. A comparison of global rating scales and detailed frequency scores as measures of specific aspects of oral L2 performance indicated that the two kinds of measurement were substantially correlated where there was sufficient variability in the data. An exploratory factor analysis of 26 variables, including measures of oral performance and academic test scores, yielded three orthogonal factors, interpreted as general English proficiency (including all the academic tests), vocabulary, and communicative style (consisting of interview variables). No separate factor was found for measures of fluency. Both the general English proficiency factor and the vocabulary factor were affected by length of residence in the L2 community, and general English proficiency was also affected by the students' age. Neither length of residence nor age was related to communicative style.

Conclusions. It was concluded that language proficiency results are strongly affected by the testing method (e.g. academic reading test, oral interview, story-telling task), and that an inherent difficulty in validating models of L2 proficiency is that measures faithfully reflecting a particular construct may not have adequate psychometric properties, while other psychometrically acceptable measures may fall short of representing the construct.

2:5 Metaphor Comprehension in Children's L1 and L2 (Final Report, Vol. I)

Purpose and design. This study compared the development of metaphor comprehension in Spanish-English bilingual children and monolingual English-speaking children, in order to test the hypothesis that metaphoric processing in bilinguals, as well as monolinguals, is constrained more by age and mental-attentional capacity than it is by language proficiency. Subjects were 20 Hispanic and 20 monolingual English-speaking children in each of three age groups: 7-8, 9-10, and 11-12 years, selected on the basis of a 'Figural Intersections Test' as being of normal mental capacity, which increased with age. An oral language proficiency test and a metaphor comprehension task in English were individually administered to each child. Hispanic children were also tested for oral proficiency in Spanish, and a subsample was tested for metaphor comprehension in Spanish. The language proficiency tests were similar to verbal IQ tests, while the metaphor comprehension task involved the oral interpretation of ambiguous metaphors, such as "my sister was a rock." The relative complexity of the children's metaphoric interpretations was coded with reference to the degree of semantic transformation involved in mapping an aspect of the vehicle (predicate) onto the topic (subject). The coding scheme had previously been shown to have good reliability and developmental validity for monolingual English-speaking children.

Findings. On the English proficiency test, Hispanic children scored significantly lower than the monolingual English-speaking children, and the Hispanic children resident in Canada for less than three years scored lower than those resident for more than three years. On the Spanish proficiency test, on the other hand, the more recent immigrants scored significantly higher than the long term residents. Performance on the metaphor comprehension task in English was, as predicted, found to be more strongly related to age and mental capacity scores than to oral language proficiency scores. While the bilingual Hispanic children did less well on the metaphor comprehension task than did the monolingual English-speaking sample as a whole, this was found to be related to the presence in the English-speaking sample of some students from a school in a higher socio-economic area. These children of middle-class background did better on the metaphor task than did the monolingual English-speaking children from the same schools as the bilingual children in working class areas. When the children of middle class background were removed from the sample, there was no main effect for language group on the metaphor scores, although the Hispanic children did less well on one of the two topics. Regression analyses indicated that the bilingual Hispanic children were similar to the subsample of English-speaking children from the middle-class neighbourhood in that English proficiency contributed little to the variance in their metaphor scores. Another finding was that conceptual structures developed in the first language appeared to facilitate metaphor comprehension in the second language, since for Hispanic children resident in Canada for less than three years, Spanish proficiency correlated more highly with metaphor scores in English than did English proficiency.

Conclusions. The findings of the study were in keeping with the hypothesis that, for bilingual as well as monolingual children, measured language proficiency was less predictive of metaphor performance than were age and non-verbal mental capacity scores. On a standardized test of English proficiency, the bilingual children scored significantly lower than their English-speaking schoolmates. On the metaphor task, however, the bilingual children performed almost as well as their English-speaking peers. This finding suggests that the metaphor task may be a more appropriate measure of conceptual skills in the second language than is a verbal IQ test.

3. CLASSROOM TREATMENT STUDIES

Several studies were undertaken to examine the relationship between instructional practices and the development of proficiency in a second language. During the first two years of the Project, a major focus was on the development and validation of a classroom observation instrument designed to capture the essential features of communication in the L2 classroom. This instrument was subsequently used in a process-product study which examined the impact on L2 proficiency of different instructional practices observed in core French classes. Two other studies grew out of the large-scale proficiency study described in 2:1 above. One of these involved the analysis of some specific aspects of language use and learning activities observed in French immersion classrooms, with a view to interpreting some of the earlier proficiency findings. The other study consisted of a classroom experiment in the French immersion setting, designed to enhance grammatical proficiency in the use of past tenses. These studies are summarized below.

3:1 Development and Validation of COLT Observation Instrument (Year 2 Report, Final Report, Vol. II)

The development of a new classroom observation scheme was motivated by the need to describe as precisely as possible some of the features of communication occurring in the second language classroom, and to distinguish between analytic and experiential orientations to language instruction. The COLT — Communicative Orientation of Language Teaching — scheme was derived from the communicative competence framework underlying the large-scale proficiency study and from a review of current issues in communicative language teaching.

Observation categories. The COLT observation scheme is divided into two parts. Part I, filled out by observers during the class, identifies different types of classroom activities and categorizes them in terms of: (a) participant organization (whole class activity, group work, individual work); (b) the content, or subject-matter, of the activity (e.g. classroom management, explicit focus on language form or function, other topics); (c) student modality (listening, speaking, reading, writing); and (d) materials in use (the type of material, length of text, intended users, and amount of control exerted on student language use). Part II of the COLT, which is later coded from a tape-recording of the class on a time-sampling basis, analyses communicative features of teacher-student interactions. Seven superordinate categories are identified: (1) use of target language (L1 or L2); (2) information gap (the level of predictability in an interaction); (3) sustained speech (length of utterances); (4) reaction to code or message; (5) incorporation of preceding utterances (how the participants react to each other's contributions); (6) discourse initiation (by teacher or student); and (7) relative restriction of linguistic form.

Validation. The observation scheme was piloted in 13 classes, mainly at the grade 7 level. There were four core French classes, two extended French and two French immersion classes, and five ESL classes in the sample. Each class was visited twice by two observers. Analysis of the Part I data entailed calculating the percentage of classroom time spent on the subcomponents of the various categories: participant organization, content, student modality, and materials. In the analysis of Part II, each verbal interaction feature was calculated as a proportion of its superordinate category. Results indicated that the COLT observation scheme was capable of capturing differences in the instructional orientation of the four types of classes. In core French

and ESL classes, for example, there was a heavier emphasis on form and more teacher control than in the extended French and immersion classes where there was a greater focus on meaning. Extended text was most often used in the immersion classes, and non-pedagogic materials were most often used in immersion and ESL classes. Sustained speech was least characteristic of the core French classes and most evident in French immersion and ESL classes. These comparative findings, intended as descriptive and not evaluative, generally met prior expectations concerning the various programs, except for some aspects of the ESL classes.

Conclusions. The ability of the COLT observation scheme to capture differences in instructional orientation was seen as an indication of its validity and as an important step toward identifying what makes one set of instructional techniques more effective than another.

3:2 The Core French Observation Study (Final Report, Vol. II)

In this process-product study, the COLT observation scheme was used to describe instructional practices in eight core French classes at the grade 11 level. Instructional differences were then analysed in relation to L2 proficiency outcomes in the different classes.

Subjects and procedures. The core French program was selected for study because the students' L2 proficiency could be assumed to derive largely from the classroom. The eight classes, from the metropolitan Toronto area, were preselected with the help of school board personnel to represent a range of L2 teaching practices. Early in their grade 11 year, the students were given a series of pre-tests of French proficiency, including some tasks from the large-scale proficiency study. The tests consisted of: (a) a multiple choice grammar test; (b) two written production tasks (a formal request letter and an informal note) which were scored for both discourse and sociolinguistic features; (c) a multiple choice listening comprehension test calling for the global comprehension of a series of recorded texts; and (d) an individual oral interview administered to a subsample of students from each class and scored for proficiency in grammar, discourse and sociolinguistics. During the school year, each class was visited four times for observation with the COLT scheme (in October, January, March and April). Observation periods lasted 40 or 70 minutes, depending on the duration of the class, and were tape-recorded. In May, the classes were post-tested with the same tests, and those students interviewed at the time of pre-testing were reinterviewed.

Analysis of COLT observations. Based on the Part I and Part II categories of the COLT observation scheme, it was possible to rank order the eight classes on a bi-polar composite scale from 'most experiential' to 'most analytic', based on the percentage of class time spent on practices defined as experiential in contrast to analytic. In the two most experiential classes, for example, there was proportionately significantly more topic control by students, more extended written text produced by the students, more sustained speech by students, more reaction (by both teacher and students) to message rather than code, more topic expansion by students, and more use of student-made materials than in the other classes. These two classes were labelled 'Type E' classes, in contrast to the remaining 'Type A' classes, where significantly more analytic features were in evidence, including a higher proportion of topic control by teachers, minimal written text by students, student utterances of minimal length, student reaction to code

rather than message, and restricted choice of linguistic items by students. The COLT analysis revealed at the same time that none of the classes was prototypically experiential or analytic, but instead intermediate along the bi-polar scale. The COLT findings were supported by teacher questionnaires providing information about classroom activities throughout the year.

The relationship of COLT findings to L2 proficiency. It was predicted that the Type A classes would be significantly higher on both written and oral grammatical accuracy measures than the Type E classes, but that the Type E classes would score higher on all other proficiency measures, including discourse and sociolinguistic measures, and scores on global listening comprehension. However, based on the post-test scores adjusted for differences in pre-test scores, no significant differences were found between the Type E and Type A classes, although a near-significant difference (p .06) emerged in favour of the Type A classes on the grammar multiple choice test. When the two Type E classes were compared to the two most analytic Type A classes (labelled Type A*), the Type A* classes did significantly better on the grammar multiple choice test (and specifically on agreement rules), but no other significant differences were found. A detailed correlational analysis relating the use of specific COLT features to L2 proficiency outcomes showed that there were both experiential and analytic features that were positively related to adjusted post-test scores. The profile of a successful core French classroom with respect to proficiency was identified as one in which a generally experiential approach was used with relatively more time devoted to features such as information gap, reaction to message, and topic incorporation. At the same time, positive correlations were found between a number of form-focussed, teacher-directed activities and proficiency outcomes. It was concluded that analytic and experiential approaches may be complementary rather than two ends of a continuum.

Qualitative analysis. In light of the fact that one of the two Type E classes made the most gain in overall proficiency and that the other Type E class made the least gain, the transcripts of these two classes were reviewed for qualitative differences that had not been captured by the COLT. The high-scoring class was found to engage frequently in communicatively rich interaction involving feedback and negotiation of meaning, while the low-scoring class received less feedback and spent more time on stereotyped routines. It was concluded on the basis of these findings that observation procedures based on COLT needed to be supplemented by more detailed forms of discourse analysis.

3:3 The Immersion Observation Study
(Final Report, Vol. III)

Classroom observations were carried out in nine grade 3 and ten grade 6 early total immersion classes in the Toronto and Ottawa regions, for the purpose of obtaining information on classroom processes and interpreting earlier immersion L2 proficiency findings. Each class was observed and tape-recorded for a full school day, including any instruction in English. All the tape-recordings were then transcribed. Analyses of some aspects of language use in the immersion classes are presented in the Project report. Further analyses are planned as time and finances permit.

Vocabulary instruction in immersion classes. L2 vocabulary-related classroom activities observed in the grade 6 immersion classes were analysed in the light of a classification scheme for describing such activities, and in relation to different kinds of linguistic knowledge involved in vocabulary learning. The analysis is qualitative rather

than quantitative. The classification scheme focusses on the following distinctions: (a) planned/unplanned instruction — the extent to which vocabulary instruction is seen as a purposeful activity; (b) systematic/haphazard instruction — the degree of systematicity with which specific features of vocabulary are taught; (c) written/oral activities — use of each medium for vocabulary instruction is seen to have a differential effect on lexical knowledge; (d) cross-linguistic/L2 based approaches to vocabulary teaching — a role for controlled use of the L1 is noted; (e) control of vocabulary selection — this may be by text authors, teacher or students; (f) the linguistic focus of teaching — whether the focus is on interpretation in context, conventional meaning, or other structural aspects of vocabulary. Linguistic aspects of vocabulary knowledge are categorized in terms of phonology, morphology, syntax, discourse semantics, graphology, and sociolinguistics. Analysis of the immersion classes in the light of these descriptors indicated that most planned vocabulary teaching occurred during reading activities organized around particular themes, during which students learned to pronounce words that they read aloud, to interpret passages, and in which the meanings of unfamiliar words were explained. Unplanned, spontaneous teaching of vocabulary was often student-initiated, as a specific word was needed to express an idea. The focus of both the planned and unplanned vocabulary teaching was mainly on interpretation and meaning. Given its association with reading activities, the teaching of new words emphasized written varieties of French and spelling. One example of how the students' prior knowledge could be drawn upon was provided by a teacher who drew attention to formal resemblances in the L2, enabling students to arrive at the meaning of an unfamiliar derived word. With some exceptions, the presentation of structural information about vocabulary was limited to the separate grammar lesson. Errors in gender, for example, tended to be only haphazardly corrected during other activities. There was no evidence that the teachers were focussing on sociolinguistic or discourse-related aspects of vocabulary. It was concluded that vocabulary teaching in the immersion classes occupied a rather narrow place in the overall teaching plan, and was mainly preoccupied with meaning and interpretation with insufficient planned attention to other aspects of vocabulary knowledge.

Vous/tu input. The underuse of vous as a politeness marker by early immersion students had been noted in the earlier proficiency study. The classroom observations provided an opportunity to relate these findings to use of vous and tu in the classroom context. Uses of tu and vous by the ten grade 6 immersion teachers and by the students in their public talk were counted and classified according to the functions they served: singular, plural, or generic; formal or informal. Teachers were found to use tu and vous about equally often, with tu generally being used to address individual children and vous to address the class as a whole. Occasionally, however, tu was used to the class and vous to individual students, leaving room for potential confusion. There was scarcely any use of vous by the teachers as a politeness marker, and its infrequency in this function in the classroom context was seen as a reason for its underuse as a politeness marker by early immersion students. Their underuse of vous in its plural function was at the same time attributed in part to the finding that, although used relatively frequently by teachers, very few opportunities appeared to arise for student production of vous plural in the classroom context. In conclusion, it was hypothesized that students would benefit from functionally-oriented instruction in the use of vous/tu, and opportunities to use them appropriately.

Student talk in teacher-fronted activities. It had previously been hypothesized that shortcomings in the grammatical competence of early immersion students may be due to

a lack of opportunity to produce 'comprehensible output', i.e. precisely conveyed messages demanding more rigorous syntactic processing than that involved in comprehension. In order to determine the opportunities that the immersion students had to talk in class, transcripts based on 90 minutes of French class time in each of the nine grade 3 and ten grade 6 classes were analysed, as well as the English portion of the day in the grade 6 classes. In general, the transcripts captured public talk rather than private, individual conversations. Each student turn was categorized according to length (minimal, phrase, clause, and sustained), and source (e.g., whether teacher- or student-initiated, preplanned or unplanned, linguistically restricted in some way or not, reading aloud from text, or reciting from memory). The findings indicated that in the 90-minute French portion of the day, student talk was less than two thirds as frequent as in the English portion of the day. Sources of student talk in French were very similar for the grade 3 and grade 6 students, the most frequent source being teacher-initiated student talk where the students' response was highly linguistically constrained, which appeared to encourage minimal responses from the students. Extended talk of a clause or more appeared to be encouraged when students initiated talk and when they had to find their own words. However, less than 15% of student turns in French were found to be sustained, i.e. more than a clause in length, when reading aloud was not included. It was concluded that greater opportunities for sustained talk in French by the immersion students are needed, and that this might be accomplished through group work, the provision of more opportunities for student-initiated talk, and through the asking of more open-ended questions by teachers.

Error treatment. An analysis of the grade 6 immersion teachers' correction of errors was based on the complete French transcripts of the ten classes observed. It focussed on the grammatical and pronunciation errors corrected by the teachers, the proportion of such errors corrected, and the systematicity of error correction. The highest proportion of error was observed in frequently used grammatical features such as gender, articles, and verbs. Only 19% of grammatical errors overall were corrected, but gender, article, and verb errors were more often corrected than other grammatical errors. About two-thirds of pronunciation errors were corrected. A lack of consistent and unambiguous teacher feedback was noted.

3:4 Functional Grammar in French Immersion
(Final Report, Vol. II)

This experimental study was designed to investigate the effect on immersion students' French proficiency of an approach to grammar teaching which involved the provision of focussed input in a problematic area of French grammar and provided students with increased opportunities for meaningful productive use of the target forms. Following a workshop with teachers, a set of classroom materials aimed at teaching the meaning distinctions between two major past tenses, the imparfait and the passé composé, were introduced for an eight-week period into grade 6 early immersion classes in six schools. These experimental classes were compared on pre-tests, immediate post-tests, and on delayed post-tests (three months later) with comparison grade 6 immersion classes in six other schools who were not exposed to the materials. The tests consisted of narrative compositions previously used in the large-scale proficiency study, as well as specially constructed cloze tests with rational deletions, and oral interviews administered to a sub-sample of students in each class. All the tests were designed to assess the students' ability to make appropriate use of past tenses and were scored

accordingly. There were two forms for each test, administered across testing sessions in a counterbalanced design.

The classroom materials. Adapted from an existing bank of activities focussing on the imparfait and the passé composé, the materials were divided into eight units, each to be used in a specific week. The teaching approach emphasized the integration of grammar teaching with worthwhile subject matter content and the personal experience of students. The oral and written activities, providing focussed input and opportunities for practice in using the two tenses, included the following: reading a simplified French-Canadian legend, discovering how the imparfait and passé composé served different functions in the legend, illustrating aspectually contrasting sentences, applying proverbs to the legend and to the students' own experiences, miming the progressive function of the imparfait, working in small groups to create new legends, and producing albums of childhood memories.

Findings. On the immediate post-tests, with adjustment made for pre-test scores, the experimental classes were significantly ahead of the comparison classes on two out of three measures: the cloze test and the oral interview. Three months later, however, at the time of delayed post-testing, there were no significant differences between the experimental and comparison groups on any of the tests. Both groups had improved their test performance over time. Evaluations of the materials by the experimental teachers at the end of eight weeks indicated general satisfaction with the materials, although some problems were noted with specific activities. Teachers indicated that they spent on average about 1 1/2 hours per week on the material. From some of their comments, it appeared that certain activities promoted more attention to subject matter content than to linguistic code, and informal observations in some classes indicated that past tense errors often went uncorrected during the 'Proverbes' activity. It was noted that one class with a teacher who was observed to provide frequent corrective feedback obtained the best results of all the classes on the composition test. Questionnaires administered to experimental and comparison group teachers at the time of the delayed post-testing indicated that the latter had also spent time working on the target verb tenses.

Conclusions. It was concluded that the teaching approach had succeeded in accelerating grammatical development in the experimental classes, but that to promote more long-term benefits some revision was needed in the materials, including more specific guidelines to teachers about the provision of corrective feedback. The fact that the control classes also appeared to have worked on past tenses was an additional factor that was surmised to have affected the long-term results.

4. SOCIAL CONTEXT AND AGE

The relationship between individual and social-environmental factors and the development of bilingual proficiency was examined in several minority and majority language learning contexts. In one large-scale study of Portuguese-Canadian students, the relationship between language use patterns, language attitudes, and bilingual proficiency was investigated by means of correlational and regression analyses, while in a small sample of beginning school-aged children of Portuguese home background, a detailed study of language interaction at home and at school was carried out with a view to relating interactional variables to later academic achievement. In another minority context, an ethnographic study focussed on students attending a French language

elementary school in Toronto. Finally, two studies examined the relationship between age and language learning: one among Japanese immigrant students of different ages and the other among Anglophone majority students learning French in three different school programs.

4:1 Language Use, Attitudes and Bilingual Proficiency of Portuguese Canadian Children (Final Report, Vol. III)

Purpose and design. In this study, the bilingual proficiency of grade 7 students from an important language minority group in Toronto was studied in relation to family background variables, the students' patterns of language use, and their language attitudes. Theoretical issues examined were: (a) the nature of language proficiency indicated by the pattern of relationships within languages; (b) the cross-lingual dimensions of language proficiency indicated by the pattern of relationships across languages; and (c) the extent to which proficiency in English and Portuguese could be predicted by language use and attitude variables.

The sample consisted of 191 students enrolled in Portuguese heritage language programs in seven inner-city Toronto schools. More than half these students were of Azorean background. The students all completed two questionnaires. One was a language use questionnaire concerning family background (e.g. birthplace, parents' language use, education, and occupations), language use patterns (use of Portuguese and English at home, in school, and in the community), and self-ratings of proficiency in English, Portuguese, and French. The other was a language attitude questionnaire which investigated dimensions such as integrative and instrumental orientations towards English and Portuguese, language use preferences in different contexts, the role of English and Portuguese in the students' ethnic identity, perceived attitudes of parents towards the students' education and language use, attitudes towards Portuguese dialects and language mixing, cultural assimilation, and attitudes towards French. Tests in English and Portuguese were also administered. In each school the students were divided randomly into three groups. One group did multiple choice grammar tests in English and Portuguese. A second group received a multiple choice discourse test in each language similar to the one administered in the large-scale proficiency study (see 2:1 above). Students in this group were also given individual oral tests in English and Portuguese, each of which contained tasks to be scored for grammar, discourse, and sociolinguistic proficiency. The sociolinguistic task in each language was adapted from the oral sociolinguistic test administered in the large-scale proficiency study. A third group of students in each school was given sociolinguistic written production tests in each language, again based on the test designed for the large-scale proficiency study.

The nature of language proficiency and its cross-lingual dimensions. A considerable degree of interrelationship was found among Portuguese self-ratings, multiple-choice discourse scores in Portuguese, and the various oral measures of Portuguese proficiency. A principal components analysis suggested a global Portuguese proficiency dimension, supplemented by academically related aspects of proficiency. Few relationships, on the other hand, were found among the measures of oral English proficiency, apparently because of a generally high level of performance giving rise to a lack of variability in scores. Across languages, self-ratings of proficiency in Portuguese, English, and French tended to be significantly related to each other. Further relatively strong cross-lingual relationships were observed for each set of written measures: i.e. between multiple choice grammar scores in English and Portuguese, between multiple

choice discourse scores in each language, and between written sociolinguistic scores in each language. These findings provided strong evidence for the interdependence of cognitive-academic skills across languages.

Predictors of bilingual proficiency development. Multiple regressions revealed that a considerable amount of the variance in the self-ratings of Portuguese proficiency could be related to attitudinal and language use variables such as students' acceptance of Portuguese, their knowledge and pride in Portuguese culture and achievements, their use of Portuguese media, exposure to Portuguese in the home, and their acceptance of and liking for French. Ceiling effects on the English self-ratings appeared to be at least partly responsible for the much weaker correlations found with attitude and use variables, although positive relationships were found with acceptance of English, use of English with siblings, and acceptance of French. The amount of exposure to Portuguese, both in Portuguese language classes and in the form of visits to Portugal, attendance at Portuguese mass, and Portuguese TV watching, appeared to be strongly related to measures of Portuguese proficiency, with weaker relationships noted between attitude variables and Portuguese proficiency. Minimal relationships were found between language use and attitude variables and the English proficiency measures, although there was evidence to suggest that positive attitudes towards Portuguese and students' use of Portuguese at home and in the community were in no way detrimental to their English proficiency.

Comparison with Azorean native speakers. A comparison of the Toronto students' test scores in Portuguese with those obtained by 69 grade 6 students in the Azores revealed that there were highly significant differences favouring the Azorean group on most measures of Portuguese proficiency. As in the large-scale proficiency study involving French immersion students (see 2:1 above), differences were most apparent on measures of grammar. The strong relationship found between Toronto students' attendance at Portuguese language classes and proficiency in Portuguese was seen as an indication that, in their minority context, more intensive exposure to Portuguese in an academic context could be advantageous for the bilingual development of the Toronto students.

4:2 Longitudinal Study of Young Portuguese Background Children: Bilingual Proficiency Development and Academic Achievement (Final Report, Vol. III)

Purpose and design. The major purpose of this ongoing study is to investigate the development of proficiency in both Portuguese and English in the transition from home to school. Twenty children from Portuguese backgrounds are being followed from the junior kindergarten year through grade 1 with respect to patterns of language interaction in the home, performance on a variety of language proficiency and literacy awareness measures, and (in grade 1) reading performance. Patterns of interaction in the home and knowledge of Portuguese and English will be used as predictors of English reading performance in grade 1. Thus, the study addresses theoretical issues such as the interdependence of L1 and L2 as well as practical issues related to the interaction between home and school variables in affecting the extent to which minority students are successful academically. The study also will provide a corpus of longitudinal data for analysis of students' developing proficiency in their two languages.

Methodology. The main sample consists of 20 Toronto students receiving the entire battery of tests. These are the Draw a Person Test, the Record of Oral Language (i.e.

sentence repetition) (English and Portuguese), Letter Identification (English and Portuguese), Concepts about Print (English and Portuguese) and, in Year 3 (Spring 1987), Test of Writing Vocabulary (English and Portuguese) and Gates McGinitie Reading Test (Comprehension subtest). (For test references, see complete study in Chapter 8.) In addition, children were taped in their homes for one and a half hours each year of the study.

Twenty-six grade 1 students (average age 7) in the Azores were also administered the Concepts about Print test, an oral interview, and Test of Writing Vocabulary (Clay 1979) in Portuguese for comparison purposes with the grade 1 Toronto data. In addition, six five-year-old students in the Azores were taped for one and a half hours in their homes. Data were also collected in Mainland Portugal from ten five-year-old children in a village situated a hundred kilometres northwest of Lisbon. A Portuguese version of the Record of Oral Language was constructed and administered to the children. Six of the ten were randomly chosen to be taped in the home.

Current status of the study. All the Year 1 home recordings have been transcribed and scoring schemes developed for grammar and pragmatics. A sample of students' transcripts have been scored but not the entire group. The Year 3 data will be collected in May and June of 1987. Subsequent to this data collection, a proposal will be developed to complete the transcription and data analysis relating home language use and proficiency in literacy-related aspects of English and Portuguese to English reading performance at the grade 1 level.

4:3 Ethnographic Study of a Toronto French Language School (Year 2 Report; see also Heller 1984)

In this ethnographic, sociolinguistic case study of a French-language elementary school in Toronto, patterns of language choice and language use were investigated in relation both to the micro-level interactional context and to the macro-level context of school and community. The study examined the role that the use of French and English played in the development of students' social identities.

Methodology. Micro-level data were collected in the school by means of participant observation over a six-month period, mainly in a grade 7/8 class, and through tape-recordings of eight students who each wore a tape-recorder for two entire school days. Four of the students were selected as ethnolinguistically representative of the school and the other four were randomly selected. Macro-level data were collected through a school-wide parent questionnaire and in interviews with school administrators, staff, members of the Parent-Teacher Association, and an ethno-linguistically representative subsample of parents.

Findings. Just over half the parents returned their questionnaires, which indicated considerable heterogeneity of family origins, linguistic backgrounds, and goals with respect to bilingualism and the maintenance of French. For example, over 40% of the families were of linguistically mixed marriages (usually with a francophone mother), 30% were francophone, 11% anglophone, and the remainder from a great variety of linguistic backgrounds. Very few parents and under half the children were Toronto-born. Family homes were widely dispersed over half of the city, making it hard for students to maintain friendships outside school. In-school observations revealed that there were three distinct groups of students: English-dominant, bilingual, and French-dominant.

The first two preferred to speak English among themselves, and the third -- a minority -- preferred French. Access to the different peer networks depended on appropriate language choice. Each group experienced its own tensions: French-dominant students reported pressure from peers to speak English outside class, while for English-dominant students, performance in French in class could be stressful. Bilingual students were observed to take part in occasional bilingual word-play and code switching, which was seen as their way of resolving the social tensions they experienced from their intermediate position and suggested that, for them, French and English were separate domains.

Conclusions. The heterogeneity of the school population and the varied linguistic experiences of the students were seen to militate against the formation of a monolithic French identity. Instead, observed patterns of language use indicated a close connection for the students between language choice and their evolving social identities.

4:4 Age on Arrival, Length of Residence, and Interdependence of Literacy Skills among Japanese Immigrant Students (Final Report, Vol. III)

Purpose and design. This study investigated the cross-lingual dimensions of language proficiency and the relationship between age and second language acquisition, with a focus on the development of reading and writing skills. We hypothesized that despite the dissimilarity of languages and writing systems, significant positive relationships would be found between Japanese minority children's L1 reading and writing skills and their acquisition of English reading and writing. An investigation of the relationships between Japanese and English proficiency appears to provide a stringent test of the interdependence hypothesis, which posits a common underlying proficiency for bilinguals, since the two languages have little in common at a surface structure level.

Subjects in the study consisted of 273 students between grades 2 and 8 attending the Japanese School of Toronto Shokokai Inc. Students were tested in May and June 1984 with measures of reading and writing in both Japanese and English. The reading comprehension subtest appropriate to students' grade level of the Gates-McGinitie Reading Test was given to all students who had been in Canada for at least six months as a measure of English reading skills. The Kyoken Standardized Diagnostic Test of Reading Comprehension published by the Research Institute for Applied Education in 1981 was given as the measure of Japanese reading skills. In addition, a letter-writing task in English and Japanese was administered to all children.

Scores on the English and Japanese reading tests were converted to T-scores to permit comparability across grades with the influence of age removed. In addition, English grade equivalent scores were used in some analyses as an approximate index of students' absolute level of English reading skills. A variety of indices of writing skills in Japanese and English were assessed.

Results. The results of correlational and regression analyses provide a consistent picture in relation to the acquisition of English reading and writing skills and their relationship to students' Japanese reading and writing proficiency. First, although the sample as a whole performs close to the mean (i.e. Japanese norms) in Japanese reading skills, there is a clear negative relationship between length of time in Canada and students' Japanese reading proficiency. The negative effect of length of residence on Japanese writing, however, appears minimal. Age of arrival in Canada appears to be a

more potent force in predicting maintenance of Japanese writing skills than length of residence. Similarly for Japanese reading, the older students are when they come to Canada, the better prospects they have for strong continued development of Japanese reading skills. This effect is not entirely due to the fact that students who arrive at older ages tend to have spent less time away from Japan, since the partial correlation between age of arrival and Japanese T-score remains significant even when length of residence is controlled.

It appears that students require about four years' length of residence, on the average, to attain grade norms in English reading skills. There appears to be some tendency for students who arrive at the age of 6-7 to make somewhat more rapid progress towards grade norms than those who arrive at older ages.

When length of residence is controlled, a significant relationship emerges between Japanese reading skills and English reading. Students' age of arrival in Canada (AOA) is also strongly related to English reading (controlling for length of residence), suggesting the influence of general cognitive maturity in mediating the cross-lingual relationship of cognitive/academic skills. General cognitive maturity, however, cannot account fully for the interdependence of reading skills across languages since significant relationships across languages were found for reading T-scores, in which the effects of age have been removed.

Writing performance was less closely related across languages than was the case for reading. This may be partly a function of the different types of measures used in each case (standardized reading tests v. non-standardized writing tasks). However, consistent significant relationships were obtained between Japanese writing and both English reading and writing measures. For some variables (e.g. Spelling) there was strong evidence of a specific cross-lingual relationship that was not mediated by more general cognitive/academic proficiencies.

Conclusions. In general, the data are consistent with previous studies in supporting the interdependence of cognitive/academic skills across languages. They also suggest that at least four years is required for students from highly educated backgrounds to attain grade norms on English academic tasks and that continued development of L1 academic skills to a high level (i.e. that of students in the home country) is a formidable task for students who arrive in the host country at an early age (particularly prior to formal schooling) but is considerably less problematic for students who arrive after several years of schooling in their home country.

4:5 Starting Age and Oral French L2 Proficiency in Three Groups of Classroom Learners (Final Report, Vol. III)

Purpose and design. The purpose of this study was to determine whether there are specific long-term advantages in oral L2 proficiency that can be associated with intensive L2 exposure at an early age in a total French immersion classroom setting. Three groups of grade 10 learners, with 11-12 subjects per group, were interviewed and given an oral sociolinguistic test in French: one group was from an early total immersion program which had begun in kindergarten, while the other two groups (from a late immersion and an extended French program respectively) had started their intensive exposure to French much later, in grade 7. A group of 12 native French speakers in grade 10 was also included in the study. The guided oral interviews were designed to

provide students with communicative contexts for the use of a range of verbs and verb forms. Transcripts of the interviews were analysed with respect to verb use and oral fluency in French. Scoring of verbs consisted of assessing the use of target verb forms in the context of specific questions, while the assessment of oral fluency was based on the nature and frequency of markers of disfluency and the linguistic contexts in which they occurred. The sociolinguistic oral test was based on the one used in the large-scale proficiency study (see 2:1 above).

Results. Group comparisons of the students' verb use indicated that the early immersion students were significantly more native-like on some variables (imparfait, conditional, use of pronoun complements in clitic position), but were no more native-like than the other learner groups on other variables such as use of number and person distinctions, time distinctions, and lexical variety, and in some instances tended to be less native-like than one or both of the other groups. The analyses of fluency revealed that in most types of disfluency, the three learner groups produced significantly more disfluencies than the native speakers but did not differ from one another. There was some evidence, however, that the early immersion students were producing fewer cut-offs and 'uh', 'um' etc. transition markers. The early immersion students were also less likely than the late immersion students to use transition markers in within-phrase locations, where such disfluencies were hypothesized to be more disruptive to discourse coherence than in between-clause or between-phrase locations. These findings indicated some advantages in oral fluency for the early immersion students who had started their intensive L2 program at a young age. Results on the sociolinguistic oral test, however, showed that the early immersion students did not manifest any general advantage over the other learner groups in sociolinguistic proficiency. While the early immersion groups displayed a slightly greater tendency to use attenuating conditional verb forms in formal social situations, they tended to be less sensitive to the appropriate use of the second person forms vous and tu than the late immersion and extended French students, whose intensive exposure to French in school had begun much later.

Conclusions. With respect to oral L2 proficiency, it appeared that there were some advantages to an early start in a French immersion program in the area of fluency and in the use of the verb system, but no advantage in the sociolinguistic domain. Some weaknesses in the verb system were also observed. As in other studies conducted in the early immersion context, a need for more emphasis on problematic areas in the target language system was indicated, along with greater opportunities for sustained oral and written expression.

5. CONCLUSIONS

The overall conclusions of the studies can be summarized with respect to the nature of bilingual proficiency and the influences on its development both in classroom and natural settings.

The nature of proficiency. At the inception of the study, the primary methodology envisaged for investigating the nature of language proficiency and its cross-lingual dimensions was confirmatory factor analysis. However, as a result of the findings of our Years 1-2 study of proficiency among French immersion students, in which little evidence emerged for the hypothesized trait structure, we became more explicitly conscious of the fact that the relationships between different components of language proficiency were a function of the specific language learning experiences to which

particular samples of individuals were exposed. This perspective implies a wider variety of analytic methods for investigating the nature of proficiency; specifically, we can discover a considerable amount about the nature of proficiency by observing its behaviour as a function of individual, social and educational conditions. Thus, we shall first consider the findings of our factor analytic studies and then examine findings of other studies that elucidate the nature of proficiency.

All studies that examined the relationships among different components of proficiency found significant correlations among written tests (including the core French observation study -- see 3:2 above). These relationships were found across languages in the grade 7 Portuguese study (4:1), the Japanese study (4:4), and the metaphor comprehension study (2:5). Some evidence emerged for an oral factor (e.g. a communicative style dimension in the "Communicative skills of young L2 learners" study -- 2:4) but the relationships among oral measures were considerably less strong than for the written measures. Similarly, some cross-lingual relationships among oral measures were found in the Portuguese grade 7 study but again the relationships were only marginally significant. These data are consistent both with the notion of a specific dimension of proficiency related to the ability to process language in context-reduced or decontextualized situations and with the hypothesis that this dimension is interdependent across languages.

There was considerably less evidence in the factor analyses for the hypothesized trait structure distinguishing grammatical, discourse, and sociolinguistic aspects of proficiency. However, discourse and grammar factors did emerge in the lexical proficiency study (2:3) and there was also some evidence for a separate vocabulary factor. The limitations of placing exclusive reliance on factor analysis for confirming hypothesized trait structures are illustrated in the fact that in this lexical study several mutually exclusive solutions produced an acceptable fit to the data. Also, in the original proficiency study (2:1), comparison of French immersion with native French speakers produced evidence that discourse skills were distinguishable from grammar and sociolinguistic skills, in that differences between L2 learners and native speakers were found only for the latter two aspects of proficiency.

Thus, consistent with the position advanced by Cziko (1983), the lack of strong support for the hypothesized trait structure in the factor analyses does not lead us to abandon the concept of traits. They are conceptually distinguishable and educationally important even if they are not statistically verifiable in relatively homogeneous school populations.

Classroom treatment. Our classroom treatment findings from different program settings lead to three main overall conclusions. First, there is evidence from both the core French and the immersion observation studies that the analytic focus and the experiential focus may be complementary rather than two ends of a continuum, and that they may provide essential support for one another in the L2 classroom. Second, the quality of instruction is clearly important in both analytic and experiential teaching. Analytic teaching will be successful in developing L2 proficiency only if it is appropriately matched to the learners' needs, while experiential teaching should involve communicatively rich interaction which offers plenty of opportunities for production as well as global comprehension on the part of the student. Third, learners may benefit if form and function are more closely linked instructionally. There is no doubt that students need to be given greater opportunities to use the target language.

Opportunities alone, however, are not sufficient. Students need to be motivated to use language accurately, appropriately, and coherently. In all these respects, the 'how' and 'when' of error correction will be a major issue for future investigation.

It seems reasonable to conclude that in all the programs under investigation — core French, heritage languages, and French immersion — much more work needs to be done in the area of curriculum design. Such work should include research to determine what combinations of analytic and experiential activities are most effective for different types of student. Another comparatively neglected area from the research point of view is teacher training and professional development. This area is likely to become more important at a time when more and more teachers are breaking away from their former dependence on prescribed pedagogic formulas and are increasingly making their own, more flexible, decisions about what can be done in the classroom.

Individual and social variables. With respect to the influence of individual and social variables on the development of proficiency, we can think of these effects in terms of the relative influence of attributes of the individual (e.g. cognition, personality) versus the target language input received by the individual. With respect to attributes, for example, it is clear from the Portuguese grade 7 and Japanese studies (4:1 and 4:4 above), as well as the immersion age study (4:5) that cognitive attributes of the learner play a significant role in at least certain aspects of target language acquisition. In the grade 7 Portuguese study and the Japanese study, children's cognitive/academic proficiency in their L1 was significantly related to the level of cognitive/academic proficiency attained in the L2. The relatively strong performance of late immersion students in comparison to those in early immersion is consistent with the notion that the learner's cognitive maturity (as indicated by age) is positively related to efficiency of L2 acquisition (at least up to the point where cognitive development reaches a plateau, possibly in the early to middle teens).

There is some evidence that cognitive attributes are more related to acquisition of certain aspects of proficiency than to others. For example, L1 cognitive/academic skills are more closely related in the Portuguese grade 7 study to performance on L2 written (context-reduced) tasks than is the case for oral tasks. Also, discourse proficiency appears to be somewhat less influenced by input/exposure variables than is the case for grammar, as illustrated by the native-speaker comparisons in the large-scale proficiency study (2:1) and Portuguese grade 7 study as well as in the regression analyses for Portuguese proficiency in the latter study (4:1).

In short, one way of thinking about the trait structure and its relationship to psychological variables is to distinguish between aspects of proficiency that are relatively more dependent on input from the environment for their full development than on attributes of the individual (e.g. oral grammar) and those that rely probably as much on individual attributes (e.g. cognitive skills, personality variables) as on input for their development (e.g. oral and written discourse, context-reduced proficiency generally). We would see sociolinguistic aspects of proficiency (particularly in the oral mode) as intermediate between grammar and discourse with respect to their relative dependence on input versus attributes. In the case of sociolinguistic proficiency, personality variables are likely to be at least as important as cognitive variables but input is clearly also crucial, as demonstrated by the immersion observation study (3:3), which showed minimal input to students regarding sociolinguistic variation. The relatively greater problems that early immersion students experience with grammar and sociolinguistic

proficiencies in comparison to discourse is consistent with this position, as is the more evident influence of exposure variables (e.g. visits to Portugal) on grammar than on discourse skills in the Portuguese grade 7 study.

In conclusion, the picture of bilingual proficiency that emerges from our studies is one of a dynamic evolving complex of traits that become differentiated from each other as a function both of variation in the input from the classroom or wider environment and the individual attributes of the learner.

Footnote

[1] In recognition that abstract, underlying language competence is not directly measurable, but inevitably coloured by the method of elicitation used, the term 'proficiency' is used in this report in a global sense to encompass both competence and performance aspects of grammar, discourse, and sociolinguistics that are measured by our tests.

References

Cziko, G. Some problems with empirically-based models of communicative competence. **Applied Linguistics**, 1983, 5, 23-38.

Heller, M. Language and ethnic identity in a Toronto French language school. **Canadian Ethnic Studies**, 1984, 16(2), 1-14.

Chapter 3

COLT OBSERVATION SCHEME: DEVELOPMENT AND VALIDATION

Patrick Allen, Maria Fröhlich and Nina Spada

1. INTRODUCTION

In recent years the development of communicative competence has become the explicit focus of many different second language programs. Although models of communicative competence (e.g. Canale and Swain 1980, Hymes 1972), and principles of communicative language teaching (Breen and Candlin 1980, Johnson and Morrow 1981, Munby 1978) have been discussed extensively in the literature and a great deal of communicative teaching material (e.g. Byrne 1977, Fletcher and Hargreaves 1980, Johnson and Morrow 1979) has been produced, very little research has been carried out to examine the relationship between actual classroom practices and the development of communicative competence.

Previous research in the area of second language teaching has often been primarily product-oriented; that is, the focus has been on examining differences in proficiency (i.e. the product) brought about by different teaching methods, e.g. the grammar translation vs. the audiolingual method (e.g. Scherer and Wertheimer 1964, Smith 1970). Although these studies had valuable aspects, they were frequently inconclusive because reference to global methods proved insufficient to distinguish between actual classroom activities; in other words, they did not take classroom processes into consideration.

To conduct a process-product study which would enable us to compare the effects of instructional differences on the development of second language proficiency, at least three prerequisites had to be fulfilled: (a) a model of communicative competence had to be posited; (b) tests to assess learners' communicative competence had to be developed; and (c) observation categories had to be created in order to relate what happens in the classroom to learning outcomes.

All three issues have been addressed in the Development of Bilingual Proficiency Project. In Year 1, a concept of proficiency was developed which proposed a componential view of communicative competence --grammatical, discourse and socio-linguistic competence (i.e. knowledge of the formal systems of lexis, morphology, syntax and phonology; knowledge of the ways in which sentences combine into cohesive and coherent sequences; and knowledge of the ways in which language is produced and understood appropriately in different contexts). The underlying hypothesis was that learners may develop competence in any of these areas relatively independently and that second language programs may differentially affect the development of these components of communicative competence. With respect to the second issue, tests were developed to measure the various competencies for immersion and core French students.

The third aspect, the need for appropriate observation categories, resulted in the development of an observation scheme referred to as COLT (Communicative Orientation of Language Teaching). The instructional variables selected for examination in the COLT scheme were motivated by a desire to describe as precisely as possible some of the features of communication which occur in second language classrooms. Our concept of 'communicative feature' was derived from current theories of communicative competence, from the literature on communicative language teaching, and from a review of recent research into first and second language acquisition. The observational categories were designed (a) to capture significant features of verbal interaction in L2 classrooms, and (b) to provide a means of comparing some aspects of classroom discourse with 'natural' language as it is used outside the classroom. One reason for undertaking this research was to investigate the claim that a knowledge of the formal aspects of language develops out of meaningful language use, rather than the other way round. According to Evelyn Hatch, "the basic assumption has been ... that one first learns to manipulate structures, that one gradually builds up a repertoire ... and then, somehow, learns to put the structures to use in discourse. We would like to consider the possibility that just the reverse happens. One learns to do conversation, one learns how to interact verbally, and out of this interaction syntactic structures are developed" (Hatch 1978: 404).

Although recent approaches to L2 instruction, e.g., communicative language teaching, emphasize the need for a more meaningful and natural use of language inside the classroom, there seems to have been little research aimed at indicating the precise differences, if any, in methodology and outcomes which distinguish these from more traditional approaches. As a result of the controversy which surrounds such ill-defined concepts as 'functional practice', 'meaningful discourse', and 'authentic language use', we decided not to attempt a definition of communicative language teaching as a general global concept, but rather to compile a list of indicators of communicative behaviour, each of which could be separately observed and quantified. We hoped that this approach would enable us to investigate the communicative orientation of L2 classrooms, especially in those cases where two or more teachers claimed to be following different pedagogic approaches.

We found that none of the existing observation instruments (e.g., Moskowitz 1970, Fanselow 1977, Naiman et al. 1978) could be adopted in its entirety for the purpose of our study. We therefore decided to develop our own observation scheme, which would contain categories to measure features of communication typical of classroom discourse, as well as categories to measure how closely these interaction patterns resemble the ways in which language is used in non-instructional settings.

2. DESCRIPTION AND RATIONALE OF THE OBSERVATION SCHEME

The COLT observation scheme is divided into two parts (see Appendix A, pp. 54-55). Part I describes classroom events at the level of activity, and Part II analyzes the communicative features of verbal exchanges between teachers and students as they occur within each activity. The decision to establish 'classroom activity' as the main unit of analysis was based on the fact that this concept is familiar to teachers and constitutes the focus around which most teaching is conceived and organized. The rationale for Part II derives from the fact that the development of communicative competence is a major concern in the current language teaching literature, and constitutes one of the basic issues in the Development of Bilingual Proficiency Project. In this section we will present a brief discussion of the main parameters of the observation scheme. The description of classroom activities will be dealt with first, followed by a presentation and discussion of the communicative features of classroom interaction proposed in this scheme.[1]

2:1 Part I: Description of Classroom Activities

Although the concept of classroom activity is intuitively and pedagogically meaningful, a clear and unambiguous theoretical definition is not easily obtained. For this reason an operational definition containing five distinct parameters was tentatively established. Each activity, including where appropriate the constituent subsections or episodes (cf. Mitchell, Parkinson, and Johnstone 1981), is described with reference to the five parameters, as follows:

I. Activity type
II. Participant organization
III. Content
IV. Student modality
V. Materials

Each parameter includes several subsections, some of which are hierarchically organized. They represent a combination of high and low inference categories. Although the parameters and their constituent categories are intended to serve a descriptive purpose, their selection is theoretically motivated in that they reflect current theories of communicative competence, and other issues in first and second language learning which have been influential in the development of L2 methodology. The five parameters of Part I are described below:

I. Activity type

The first parameter of the observation scheme is open-ended, that is, no predetermined descriptors have to be checked off by the observer. Instead each activity is separately described: e.g. drill, translation, singing, discussion, game, dictation, role-play, reading aloud. Frequently, activities consist of two or more episodes: e.g. (a) the teacher reads the words of a song aloud, (b) the students repeat the words after the teacher, (c) the students sing the song. These would be described as three separate episodes within one activity. The parameter 'activity type' was left open so that the scheme could accommodate the wide variety of activities occurring in various L2 programs at different age levels.

II. Participant organization

This parameter describes three basic patterns of organization for classroom interactions: Is the teacher working with the whole class or not? Are the students divided into groups or are they engaged in individual seat work? If they are engaged in group work, how is it organized? The various subsections are as follows:

1. Whole class

   (a) Teacher to student or class, and vice versa (one central activity led by the teacher is going on; the teacher interacts with the whole class and/or with individual students).

(b) Student to student, or student to class and vice versa (Students talk to each other, either as part of the lesson or as informal socializing; one central activity led by a student may be going on, e.g. a group of students act out a skit and the rest of the class is the audience).

(c) Choral work by students (The whole class or groups participate in the choral work, repeating a model provided by the textbook or teacher).

2. Group work

(a) Groups all work on the same task.

(b) Groups work on different tasks.

(Note: If possible, we indicate the number of groups and the number of students in each group. We also indicate whether the teacher or the students specify the activities and the procedures, and the extent to which the teacher monitors group work).

3. Group and individual work

(a) Individual seat work (Students work on their own, all on the same task or on different tasks).

(b) Group/individual work (Some students are involved in group work, others work on their own).

The above low-inference categories are descriptive of how the students are organized as participants in classroom interaction; however, the categories may also reflect different theoretical approaches to teaching. In the literature on communicative language teaching, for example, group work is considered to be an important factor in the development of 'fluency skills', or communicative competence (Brumfit 1981; Long, Leslie, McLean, and Castanos 1976). The reason for this claim is that teacher-centred approaches are thought to impose restrictions on the growth of students' productive ability. In classes dominated by the teacher, students spend most of their time responding to questions and rarely initiate speech. Moreover, student talk in teacher-centred classrooms is frequently limited to the production of isolated sentences which are assessed for their grammatical accuracy rather than for their communicative appropriateness or value. Because the emphasis in group interactions is more likely to be on the expression of meaning, and less likely to be on the linguistic accuracy of utterances, classes which can be shown to provide more group activities may affect the L2 development of learners in ways which are different from those that represent a teacher-centred 'lock-step' approach to instruction.

III. Content

The 'content' parameter describes the subject-matter of the activities; that is, what the teacher and the students are talking, reading, or writing about or what they are listening to. Three major content areas have been differentiated: Management, Language, and Other topics. The rationale for these categories arises from current discussions of theoretical issues in first and second language acquisition, including theories of communicative competence, and also from a number of practical pedagogic concerns. The content categories are as follows:

1. Management

   (a) Classroom procedures
   (b) Disciplinary routines

2. Explicit focus on language

   (a) Form
   (b) Function
   (c) Discourse
   (d) Sociolinguistics

3. Other topics

   (a) Narrow range of reference
   (b) Limited range of reference
   (c) Broad range of reference

4. Topic control

   (a) Control by teacher
   (b) Control shared by teacher and student
   (c) Control by student

The first content category, Management, has been separated from the other content areas because it does not fall within the range of planned curriculum content, but arises from the needs of the classroom situation. Management exchanges are of particular interest in L2 learning because they often include examples of spontaneous communication within the context of an otherwise grammatically-oriented classroom (Brumfit 1976, Long 1983). Management also relates to authentic communication in that the giving and receiving of directives of a procedural or disciplinary nature represents an aspect of language use which is very common in the 'real world' outside the classroom.

The content areas Language and Other topics reflect the distinction between first language acquisition in natural settings, and second language learning in the classroom. It has been repeatedly shown that in interactions with children acquiring their first language the focus is on the message being conveyed, and that the vast majority of corrections by caretakers refer to violations of meaning rather than of form (see Snow and Ferguson 1977 for a discussion of this issue). The focus in the L2 classroom, however, has typically been on the presentation of the language code and on the correction of formal errors, especially in programs based on the grammar-translation or the audiolingual approaches. In view of the often limited success of more traditional methods of L2 teaching and the claim that the process of L2 learning is in many ways similar to that of first language acquisition (Corder 1971, Richards 1973), it has been argued that L2 teaching methods should attempt to approximate the conditions under which young children learn their first language. The question of whether the primary focus of instruction should be on meaning or on code is one of the crucial issues in this debate.

'Explicit focus on language' and 'Other topics' are both divided into several subsections. With regard to explicit focus on language, 'form' refers to grammar, vocabulary, and pronunciation, 'function' to illocutionary acts such as requesting, apologizing, and explaining, 'discourse' to the way sentences combine into cohesive and coherent sequences, and 'sociolinguistics' to the features of utterances which make them appropriate to particular social contexts. These four categories have been derived from theories of communicative competence reflected in the work of Hymes (1972), Morrow (1977), Munby (1978), Wilkins (1976), Canale and Swain (1980) and others, and on the model of L2 proficiency proposed in the Year 1 Report of the Development of Bilingual Proficiency Project. The assumption underlying the Language categories is that instruction which gives differential attention to these areas of competence may affect language learning in a variety of ways.

With respect to Other topics, an attempt was made to find a small number of superordinate categories to represent the potentially vast number of topics which can arise in conversation. We tentatively suggest a tripartite system, i.e., topics of narrow, limited, and broad range of reference. Underlying this classification is a belief that the cognitive content of instruction may have an effect on L2 learning. Topics of narrow range refer to the immediate classroom environment, and to stereotyped exchanges such as 'Good morning' or 'How are you?' which have phatic value but little conceptual content. Included in this category are routine classroom references like establishing the date, day of the week, what kind of weather it is, etc., or the use of other information which is easily verifiable or recalled. Topics of limited range refer to information which goes slightly beyond the classroom while remaining conceptually limited. Examples would be routine social topics like movies, hobbies, and holidays; school topics including extracurricular activities; and topics which relate to the students' immediate personal and family affairs. Topics of broad range go well beyond the classroom and immediate family environment, and involve reference to controversial public issues, current world events, abstract ideas, and reflective personal information such as 'What do you like about living in Toronto?' It is often the case that when such topics are under discussion ideas do not come automatically but require some degree of soul-searching and originality. Communicative theorists believe that more time should be spent promoting realistic broad-range discussions in the L2 classroom, rather than confining students to the predictable routines of model dialogues and structural drills.

The final category relating to content is Topic control, that is, who selects the topic that is being talked about: the teacher, the student, or both? Second language programs differ widely with regard to the behaviours included in this category. It has frequently been pointed out, for example, that the audiolingual method constitutes a strong claim about the role of the teacher in L2 education. In the literature on communicative language teaching, on the other hand, the teacher is not seen as an authority figure or director of the student's work, but more as a counsellor, resource person and guide. In a communicative curriculum such as the one proposed by Breen and

Candlin (1980) the teacher and the students are seen as 'co-participants' and 'joint negotiators' of the teaching process, and the students actively participate in the selection of materials, topics and tasks. It was hoped that a close observation of classes which differ in terms of topic control, together with an analysis of classroom treatment and learning outcomes, would enable us to throw some light on the question of what constitutes the most effective balance between teacher and student roles in L2 education.

IV. Student modality

This section identifies the various skills which may be involved in a classroom activity. The focus is on the students, and the purpose is to discover whether they are listening, speaking, reading, or writing, or whether these skills are occurring in combination. A category 'other' is included to cover such activities as drawing, modelling, acting, or arranging classroom displays. We anticipated that a differential focus on the various skills and their combinations might directly affect the development of particular aspects of the learner's L2 competence.

V. Materials

This parameter introduces categories to describe the materials used in connection with classroom activities. In addition to the type of materials involved (written, audio, visual) consideration is given to the original source or purpose of the materials, and to the way in which they are used. In the case of written or audio texts, we note whether they are minimal in length (captions, isolated sentences, word lists) or extended (stories, dialogues, connected paragraphs). The categories for materials were as follows in the development study:

1. Type of materials

    (a) Text
    (b) Audio
    (c) Visual

2. Length of text

    (a) Minimal
    (b) Extended

3. Source/purpose of materials

    (a) Pedagogic
    (b) Semi-pedagogic
    (c) Non-pedagogic

4. Use of materials

    (a) Highly controlled
    (b) Semi-controlled
    (c) Minimally controlled

The third category involves us in making a judgement about whether the materials were specifically designed for L2 teaching (i.e., pedagogic), or whether they were originally intended for some other purpose (non-pedagogic). Frequently, materials from outside the school environment are adapted for instructional purposes, hence the need for an intermediate category. A real newspaper or magazine used in the classroom in its original form would be an example of real-world, non-pedagogic, or 'other purpose' material. On the other hand, a simplified reader, or a textbook unit contrived to illustrate a particular grammatical point, would be an example of materials specifically designed to be used for L2 instruction. In between, there is a category of semi-pedagogic material which utilizes real-life objects and texts, but in a modified or simulated form. An example of this might be a series of pictures or headlines from real newspapers, presented in a textbook with accompanying captions and exercises, which make the material more appropriate for the needs of the L2 learner. Advocates of the communicative approach have claimed that 'authentic' materials are essential in order to prepare students for the kinds of discourse they will encounter outside the classroom (Breen 1982, Brumfit' 1981, Phillips and Shettlesworth 1975). One of the questions we wanted to investigate was the way in which classrooms actually differ in the repertoire of materials used, and how the differences may affect the type of L2 abilities that students acquire.

In the development study the final Part I category referred to the way in which the materials are used, as distinct from the type of materials they are. The use of materials in the classroom may be highly controlled, semi-controlled, or minimally controlled. For example, consider three situations in which students are being asked comprehension questions based on a reading passage or picture. In the first situation the discourse may be highly controlled in that the questions and answers adhere quite closely to the text. In the second situation the discourse is semi-controlled, i.e., it extends occasionally beyond the restrictions imposed by the textbook. In the third situation the textbook simply provides the starting-point, and the ensuing conversation ranges widely over a number of topics which emerge spontaneously from the contributions of the students. It has been suggested, as a general principle, that a flexible treatment of materials, particularly texts, will enable students to develop their fluency, to "do many things which are not entirely predictable ... but which will indicate that their natural language learning capacities are being exercised and encouraged" (Brumfit 1981: 48).

2:2 Part II: Communicative Features

The second part of the COLT observation scheme consists of an analysis of the communicative features occurring within each activity. As in the case of the categories of Part I, the communicative features were motivated by numerous discussions in the current literature concerning communicative competence, communicative language teaching, and first and second language acquisition. The following seven communicative features have been isolated:

I. Use of target language
II. Information gap
III. Sustained speech
IV. Reaction to code or message
V. Incorporation of preceding utterances
VI. Discourse initiation
VII. Relative restriction of linguistic form

All the features are coded for teachers and students, with the exception of discourse initiation and relative restriction of linguistic form, which are coded for students only. A discussion of the seven features follows.

I. Use of target language

This communicative feature is designed to measure the extent to which the target language is used in the classroom. It is based upon the obvious assumption — not necessarily evident in all teaching methods -- that in order for a second language to be acquired it must be used by the students. This feature is covered by two categories in the coding scheme: 'Ll' refers to use of the first language, and 'L2' refers to use of the second, or target, language.

II. Information gap

This communicative feature refers to the extent to which the information requested and/or exchanged is unpredictable, i.e., not known in advance. Theories about the nature of communication emphasize that a high degree of unpredictability is characteristic of natural language use (Breen and Candlin 1980, Morrow 1981, Widdowson 1978, Canale 1983). In other words, communication must have a purpose — the giving, receiving, or requesting of information. It is not surprising that if the information requested is already known in advance, as is often the case in L2 classrooms, the motivation to communicate tends to be rather weak.

Although studies of first language acquisition have shown that there is a high level of predictability in many interactions between caretakers and children in the early stages (MacLure and French 1981), the information gap increases rapidly as language proficiency develops. In contrast, it appears that many L2 classroom interactions, even at the intermediate and advanced levels, are marked by an absence of real information gap. Students may perceive very little reason to listen carefully or to think about what they are saying when the main purpose of the exercise is to display their knowledge of grammar without consideration of the message being conveyed (cf. Mehan 1979). It follows, then, that one of the aims of communicative language teaching is to engage learners in activities where the message is reasonably unpredictable, in order to develop information processing skills in the target language from the earliest possible stage (cf. Johnson 1982).

The categories designed to capture this feature in the development study were the following:

1. Requesting information

   (a) Pseudo-requests (The speaker already possesses the information requested).

   (b) Genuine requests (The information requested is not known in advance).

2. Giving information

(a) Relatively predictable (The message is easily anticipated in that there is a very limited range of information that can be given. In the case of responses, only one answer is possible semantically, although there may be different correct grammatical realizations).

(b) Relatively unpredictable (The message is not easily anticipated in that there is a wide range of information that can be given. If a number of responses are possible, they provide different information).

III. Sustained speech

This communicative feature is intended to measure the extent to which speakers engage in extended discourse, or restrict their utterances to a minimal length of one sentence, clause or word. The rationale for this feature is primarily pedagogic. Although communication outside the classroom consists of minimal as well as sustained discourse, L2 classrooms often restrict the length of the learner's output to one sentence or less, and rarely provide opportunities for more extended speech (McEwen 1976, Bialystok et al. 1979, Mitchell et al. 1981). If practice with normally sustained discourse is considered to be important for the development of fluent speaking and listening skills, then it is necessary for the teacher to create situations where such practice can take place. The categories designed to measure this feature are:

1. Ultra-minimal (utterances which consist of one word — coded for student speech only).
2. Minimal (utterances which consist of one clause or sentence — for the teacher, one-word utterances are coded as minimal).
3. Sustained speech (utterances which are longer than one sentence, or which consist of at least two main clauses).

IV. Reaction to code or message

The fourth feature coded in Part II is closely related to the 'content' parameter of Part I — the point at issue being whether the purpose of an exchange is to focus on the language code (i.e., grammatical correctness) or on the message, or meaning, being conveyed. Research has shown that in first language acquisition attention is focussed on the meaning rather than on the well-formedness of utterances (Snow and Ferguson 1977, de Villiers and de Villiers 1979, Wells 1981). Moreover, it appears that when children are acquiring their first language, correction of the code tends to confuse rather than help the learner (Brown 1980, McNeill 1966). In the L2 literature, it has been suggested that greater opportunities to focus on meaning will help the learner approximate first language acquisition conditions, and may lead to similar success (Macnamara 1973). In the development study, this feature was covered by a single category, 'Explicit code reaction', defined as 'A correction or other explicit statement which draws attention to the linguistic incorrectness of an utterance'.

V. **Incorporation of preceding utterances**

In conversation there are many ways in which participants may react to each other's contributions. One person may add a comment, or elaborate on a preceding utterance. Another may ask a related question, or perhaps there may be no reaction at all. Some studies of first language acquisition have suggested that expansions of a child's utterance which add or request additional information and in which somewhat novel forms are used tend to enhance the development of the child's linguistic competence (Cross 1978, de Villiers and de Villiers 1979, Ellis and Wells 1980, Wells, Montgomery, and MacLure 1979, Wells 1981). Generally speaking, these studies suggest that "the best environment for learning language contains a rich variety of sentences closely tied to what the child currently produces" (de Villiers and de Villiers 1979: 109). It seems reasonable to suppose that the same principle may apply in L2 learning.

To allow coding for a limited selection of reactions to preceding utterances, six categories were included in the development study. These were ordered according to their potential for stimulating further topic-related discourse, as follows:

1. No incorporation: No feedback or reaction is given.
2. Repetition: Full or partial repetition of previous utterance/s.
3. Paraphrase: Completion and/or reformulation of previous utterance/s.
4. Comment: Positive or negative comment (not correction) on previous utterance/s.
5. Expansion: Extension of the content of preceding utterance/s through the addition of related information.
6. Elaboration: Requests for further information related to the subject matter of the preceding utterance/s.

VI. Discourse initiation

In first language communication among adult speakers as well as between children and adults, interactants generally have equality in discourse roles and rights, that is, they may not only respond to elicitations but they may also spontaneously initiate talk. From an early age, children begin to engage in complex patterns of turn-taking behaviour. It has been noted that in many mother/child interactions it is the child who initiates the exchanges, and the mother — the 'teacher' as it were — who responds (cf. MacLure and French 1981). These self-initiations are a gamble on the part of the child, an exploration of different linguistic means to negotiate meaning. Thus children create an opportunity to test their own hypotheses about the language by forcing their interactants to provide them with feedback and further input.

In many L2 classrooms the discourse roles of the learners seem to be the reverse of their counterparts outside the classroom. The classroom appears to be an environment which requires far more elicited than self-initiated talk, thus restricting the purposes for which language can be used. It follows that another principle of communicative language teaching is that students should be encouraged to initiate discourse themselves, instead of always having the role of respondent to questions imposed on them. To measure the frequency of self-initiated turns by students in different types of classroom, the category 'Discourse initiation' was included in the coding scheme.

VII. Relative restriction of linguistic form

In mother tongue communication speakers use a wide variety of linguistic forms to express the meanings they wish to convey. Apart from sociolinguistic constraints imposed, for example, by the situation or by the relative status of the interactants, the grammatical structures and semantic choices are virtually unrestricted. The same lack of restriction is evident in the speech of children acquiring their first language. As indicated earlier, children experiment with language, try out their own strategies for communication and — as their systematic errors reveal — develop and test hypotheses about the language being learned. This constant process of meaning negotiation and hypothesis testing appears to be a crucial factor in first language acquisition.

By contrast, L2 learners are typically expected to mimic specific grammatical patterns in repetition or substition drills, and are rarely encouraged to experiment or to use language freely. Often the fear is that creative, uncontrolled language use will lead to many errors which might then prove difficult to eradicate. The literature on communicative language teaching emphasizes the need for activities in which learners can practise getting a message across with whatever resources happen to be available, thus developing the type of skill which is referred to as 'strategic competence' (Canale and Swain 1980). As in mother tongue acquisition errors are viewed positively, and are considered to be a necessary step in the active process of hypothesis formation and gradual approximation to the target language: "The student must be allowed to grope, to play around with the language, to internalize it by using it and in using it to make mistakes" (Brumfit 1981: 49). As with all the communicative features, however, it remains an empirical question what techniques are pedagogically most effective in a given classroom.

In the development study, three subcategories were proposed to permit an investigation of the effect of different degrees of restriction on the development of L2 proficiency:

1. Restricted use: The production or manipulation of one specific form is expected, as in a transformation or substitution drill.

2. Limited restriction: There is a choice of more than one linguistic form but the range is very narrow, e.g. responses to Yes/No questions, statements about the date, time of day, etc.

3. Unrestricted use: There is no expectation of any particular linguistic form, as in free conversation, oral reports, or personal diary writing.

2:3 Coding Procedures

Two sets of coding procedures were developed: one for the activity level analysis (Part I) and one for the exchange level analysis (Part II).

All coding in Part I is done in real time by two observers who are present in the classroom during the observation period. The activities are timed, and the starting time for each activity is entered in the left-hand margin of the coding form. In addition to a written description of the type of activity (e.g. drill, dialogue repetition, conversation, etc.), the observers place a check mark in the appropriate boxes under each of the four major headings: participant organization, content, student modality, and materials. In the course of a single activity, several subsections may be marked. For example, under the category 'participant organization' there may be instances of student-to-student interaction, teacher-to-student interaction, and teacher-to-class interaction. In cases like this, check marks are placed in the appropriate boxes for each of these participant interaction types, and a circle is drawn round the check mark in the box which represents the primary focus or predominant feature of the activity. This procedure is followed when coding all the Part I categories.

Part II coding is performed subsequent to the lesson, and is based on an audio-recording of each of the classes observed. A time-sampling procedure within activity types is followed. Coding starts at the beginning of each activity for one minute and is resumed after a two-minute interval. During the one-minute coding periods, the frequency of occurrence of each subcategory of the communicative features is recorded by two coders. For an example of how the coding is performed, consider the following interaction between a teacher and two students which occurred within a one-minute coding period:

	Utterance	Communicative features
T:	What's the date today?	L2/pseudo-request/minimal speech
S_1:	April 15th.	L2/predictable information/ultra-minimal speech/limited form
T:	Good.	L2/comment/minimal speech

T:	What's the date today?	L2/pseudo-request/minimal speech
S_2:	April 15th.	L2/predictable information/ultra-minimal speech/limited form
T:	Good.	L2/comment/minimal speech

Consider now the following interaction between a teacher and a student which was coded somewhat differently:

	Utterance	Communicative features
T:	What did you do on the weekend?	L2/genuine request/minimal speech
S:	I went to see a movie.	L2/ giving unpredictable information/ minimal speech/unrestricted form
T:	That's interesting. What did you see?	L2/comment/elaboration (genuine request for information)/sustained speech
S:	E.T. I really liked it. He's so cute.	L2/giving unpredictable information/sustained speech/ unrestricted form
T:	Yes, I saw it too and really liked it. Did anyone else see it?	L2/comment/expansion/elaboration (genuine request for information)/ sustained speech.

It will readily be seen that the first example represents a stereotyped routine marked by pseudo-requests, predictable responses, and minimal speech patterns, while the second is much closer to natural language behaviour, and includes genuine requests, unpredictable responses, and a reasonable amount of sustained speech.

The intention is that the coding procedures for Part I and Part II should permit the investigators to provide a detailed description of the type of activities that are taking place in L2 classes, together with a characterization of these activities in terms of a wide range of linguistic-communicative and pedagogic factors that are thought to influence L2 learning.

3. THE VALIDATION STUDY

3:1 Design of the Study

The aim of the validation study was to pilot-test the COLT observation scheme in a variety of instructional settings. It is important to emphasize that the study was not intended to evaluate the second language classes and programs observed, but rather to determine whether this particular observation scheme was capable of capturing differences in the communicative orientation of L2 classrooms.

Sample. The study was conducted with a total of 13 classes, predominately at the grade 7 level, in four different second language programs: four core French classes, one history and one language arts class in both the extended French and French immersion programs, and five ESL classes.

On a weekly basis, the core French classes received an average of 129 minutes of instruction, the extended French class an average of 413 minutes, and the immersion classes approximately 800 minutes. In the ESL classes the weekly amount of instruction varied from 900 to 1800 minutes. these classes were 'self-contained'; i.e., students spent all or most of the day with the ESL teacher. In addition to English language instruction, students also received varying amounts of subject matter instruction.

The students in the core and extended French classes had started to learn French at different grade levels, ranging from grades 1 to 5. For the immersion students the length of exposure to French in school was more homogeneous. All of them had started with full immersion in senior kindergarten. There may have been some differences in the later grades regarding the proportions of the day allotted to French and English instruction respectively; however, no information was obtained on this aspect.

As was expected, the amount of previous ESL instruction varied greatly within each ESL class. Some students had arrived a few months prior to our observation, others had been in ESL programs for three years.

Regarding the students' language background, the FSL classes consisted predominantly of monolingual anglophones (96.44%). Of the ESL students, the majority were of Chinese or Vietnamese origin (70.25%), with students speaking English as a second dialect constituting the second largest group (10.74%).

The study was begun with a number of tentative expectations about the main characteristics of the four types of program. These expectations were based on some preliminary classroom observations, discussions with teachers, consultants, and school board officials, and a review of textbooks and other teaching materials. Core French is taught as a subject within a limited time period, and classes in this program were expected to contain a relatively high proportion of form-focused, teacher-centered activities. Since extended French involves the presentation and discussion of subject matter material in addition to core French instruction, the language teaching in this program was expected to be somewhat less structured and more meaning-oriented. French immersion is designed for students to receive the same education as they would in the regular English program, except that the medium of instruction is French; French immersion classes were therefore expected to provide greater opportunity for authentic discourse and for the negotiation of significant meaning. ESL teaching in Toronto differs from the three types of French program, since many more opportunities for English language acquisition exist outside the classroom. As a result, it was expected that ESL teachers would tend to use class time to practice various aspects of the language code but that they would also seek to introduce communicative enrichment material from the "real world" outside the classroom whenever possible.

It should be emphasized again that the purpose of the study was to validate the observation instrument, rather than to evaluate the programs described above. The reason for including classes from different L2 programs was to ensure that the COLT categories were capable of describing the activities in a range of instructional settings. No claim was being made at this stage in the research that the classes selected constituted a representative sample from each program.

Procedures. Each class was visited twice by two observers. The observation period per visit varied from 30 to 100 minutes, depending on the length of the lesson. The classes were recorded on audiotape, with one exception. In this case, instruction was totally individualised, i.e., the students were working on different topics, with different teaching materials, for varying lengths of time. For this reason, it was not possible to audiotape the class. Furthermore, only one of the Part I categories, Participant organisation, could be reliably observed. Therefore, with the exception of Table 1, the tables in this section report on the results of twelve classes only.

The coding procedure was similar to the one used in the development study. Although the coding of Part I and Part II was carried out independently, the coders checked their entries for Part I immediately after each observation period and, for Part II, after each minute of coding. Wherever necessary in coding Part II, the tape was replayed and any problems were discussed. For this reason, it was not considered necessary to calculate intercoder reliability coefficients.

3:2 Analysis and Findings: Part I

Initial analysis of the Part I data consisted of calculating the percentage of classroom time spent on individual categories under each of the four major headings. These calculations were carried out separately for each visit. Subsequently, tables were prepared to present the average percentage of observation time coded for various categories by class and by program.

To illustrate, let us consider two hypothetical 30-minute visits to Class 1 and Class 2 in Program X. During the first visit, Class 1 spent 10 minutes in group work, and for the remaining 20 minutes the teacher interacted with the whole class (T -- S/C). During the second visit, the class spent 15 minutes in whole class interaction, and 15 minutes in group work. In Class 2, T -- S/C interaction was coded as the dominant activity for the whole class time during both visits. The following calculations were carried out:

Participant organization - percentage of time by visit

	Whole class T -- S/C		Group	
	Visit 1	Visit 2	Visit 1	Visit 2
Class 1	66.67	50.00	33.33	50.00
Class 2	100.00	100.00	0	0

Participant organization - percentage of time by class

	Whole class T -- S/C	Group
Class 1	58.34	41.66
Class 2	100.00	0

Participant organization - percentage of time by program

Whole class T -- S/C	Group
79.17	20.83

It should be noted that the primary category checked off during an activity or the one which occurred exclusively always received credit for the entire length of time that the activity lasted. During an activity in which the teacher and students were interacting meaningfully, for example, the occasional choral repetition of a word or phrase would not be recorded. Therefore, only those categories which were recorded as the exclusive or primary features of an activity will be presented in the following tables.

Participant organization. For the first major heading, percentages were calculated for the following categories: Whole class, Group work, Individual seat work, and Combination of group work/individual seatwork. Whole class is further subdivided as follows: Teacher interacting with the whole class or individual students, Students interacting with class or individual students while one central activity is going on, and Choral work. The mean percentages by program are shown in Table 1 (p. 50).

As previously indicated, the study was begun with various expectations about which categories would best describe the four types of programs. In core French, the expectation was that there would be a great deal of whole class interaction with the teacher addressing either the whole class or individual students, as well as a substantial amount of choral work. Whole class interaction, but not choral work, was thought to be characteristic of extended French and French immersion programs. In the ESL classes, more group work than whole class interaction was expected.

The data support these expectations to the extent that all the FSL programs were characterized by a considerable amount of whole class interaction. However, the expectation about choral work in core French was not supported, since the core French mean of 14.4% for Choral work was largely attributable to one particular class. In the ESL classes observed, individual seat work -- and not group work, as expected -- predominated.

Content. For the second major heading, percentages were calculated for the following categories: Management, Explicit focus on language, Other topics, and Topic control, each of which is further divided into a number of subcategories (see section 2:1 above). For these categories, the expectation was that there would be predominant focus on form in core French; focus on form as well as other topics (particularly of limited and broad range) in extended French, and relatively greater focus on meaning than form in French immersion and ESL.

Percentages were calculated first for those categories which had occurred exclusively or had been marked as the primary feature of an activity. For example, during one activity, a teacher may have focused exclusively on grammar (Form). During another activity, Form and Sociolinguistics may have been checked off, but because the teacher had made only a brief reference to sociolinguistic aspects of language use, Form was considered the primary focus. Percentages were then determined for those categories which had occurred in combination: that is, in situations in which the observers felt that two categories had received roughly equal emphasis. For example, a core French class was practising verb endings for the second person singular and plural, with explicit reference to the difference in the use of tu/vous when addressing friends and strangers. Thus, the content of this activity was simultaneously Form and Sociolinguistics.

The mean percentages of total observed time for Content categories are presented in Table 2 (p. 50). In comparing programs, it becomes apparent that in the core French and ESL classes, more than half of the class time observed involved activities which focused exclusively or primarily on Form (58.44% and 66.43%, respectively). By contrast, in the extended French and the French immersion program, the focus on Form decreases and the focus on meaning (i.e., Other topics) increases (40.58% and 62.53% in extended French and French immersion, respectively).

This shift can be largely attributed to the teaching of subject matter, since subject matter was coded as Other topics-broad range of reference. It is interesting to note that the extended French program occupies something of a middle position between core French and French immersion; Form is given substantial weight (slightly more than one third of the observation time, if combinations are included), although considerably less than in core French and considerably more than in French immersion. It should be pointed out, however, that the difference in the emphasis on Form between French immersion and extended French may be attributable in part to the lower proficiency level of the students in the latter program; at the time of the observations, students had been in the extended French program for only a few weeks.

It is also important to note which categories of Content were seldom or never coded. One example is Discourse, which was defined as "explicit focus on the way sentences combine into cohesive and coherent sequences." Although students were exposed to oral and written discourse through listening and reading activities, explicit reference to aspects of cohesion or coherence was never made. Another category which rarely appeared in the classes observed was Sociolinguistics. The major exception was one of the French immersion classes, in which the language appropriate for journalistic reports and advertisements was compared and discussed during an entire lesson.

In summary, with the exception of the ESL classes and their unexpectedly strong emphasis on Form, the data supported initial expectations.

Topic control. The last set of Content categories reflect Topic control: that is, who selects the topic and controls what is being read, written, or talked about.[2] The data for these categories are presented in Table 3 (p. 50). As expected, teachers controlled topic selection and content most of the time in all four programs. Again, core French and ESL, which have the two highest percentages of teacher control and the two lowest percentages of teacher/student control, appear to be most similar.

Student modality. The data for Student modality — i.e., the particular skill or combination of skills involved in a classroom activity — are presented in Table 4 (p. 51). Although these categories present useful information about the amount of time devoted to listening, speaking, reading, and writing, they provide no insight into how these skills were being developed. Thus, the parameter of Student modality does not directly address the issue of whether skills practice was more communicatively-based in one program than another. In the COLT, such differences would have to be captured in the open-ended description under Activity.

Materials. The final major heading in Part A of COLT is Materials. In this report, differences among the programs in Type and Source of materials are presented.[3]

Type: Materials were classified as Text, Audio, or Visual. Since the development of discourse competence may be affected by the extent to which students are exposed to extended written texts rather than to isolated, disconnected sentences, Text was subdivided into Minimal and Extended. Mean percentages by program are presented in Table 5 (p. 51), which shows that Text was used predominantly in all programs and that Visual played a substantial role only in core French.

It was expected that the use and production of minimal text would predominate in core French, that a balance between minimal and extended text would be found in ESL and extended French, and that extended text would predominate in French immersion. These expectations were based on the assumption that classes which focused more on teaching the language code would likely include more activities involving minimal texts (e.g., worksheets with grammatical exercises) than would programs which incorporated subject matter instruction. With the exception of ESL, the data supported these expectations.

Source, the second subcategory of Materials, refers to the origin and purpose of the teaching materials used. Were the materials designed for L2 teaching and learning (i.e., Pedagogic), or were they originally intended for some other purpose (i.e., Non-pedagogic)? A third possibility is that non-pedagogic, or 'authentic,' materials may have been adapted for instructional purposes, in which case they would be coded as Semi-pedagogic.

Table 6 (p. 51) presents data on the origin/purpose of teaching materials by program. Pedagogic materials comprised the largest percentage across all programs. They were used most extensively in core French (83.69%), followed by extended French (72.88 %), French immersion (67.56%), and ESL (63.99%).[4] Non-pedagogic materials were used relatively frequently in the French immersion and ESL settings (24.13% and 15.75%, respectively), but rarely in the other two programs.

3:3 Analysis and Findings: Part II

Part II of the COLT observation scheme analyzes the communicative features of verbal interaction during classroom activities. As already indicated (section 2:2 above), it consists of the following seven communicative features: Use of target language, Information gap, Sustained speech, Reaction to message/code, Incorporation of preceding utterances, Discourse initiation, and Non-restriction of linguistic form. All of these categories are used for coding teacher and student talk, with the exception of Discourse initiation and Restriction of linguistic form, which are used for coding student talk only.

To compare communicative features of verbal interaction across programs, each category in Part II was calculated as a proportion of its superordinate feature. For example, in the core French program the proportion of L2 use within the superordinate category Use of target language was .96; the proportion of L1 use was .04. These proportions are presented by program in bar graphs; Figures 1 and 2 present the data for teacher and student verbal interaction, respectively[5]. In addition to a descriptive comparison, One-way Analyses of Variance and Duncan Multiple Range Tests were conducted to examine if the differences between programs reached statistical significance.

Teacher verbal interaction. As indicated in Figure 1 (p. 52), teachers used the target language most of the time in all four programs. They generally gave unpredictable information: for example, by giving directives or presenting new information. No significant differences between programs were found for these categories.

Teachers did not generally ask genuine questions -- i.e., questions to which they did not already have the answer. Although differences between programs did not reach statistical significance, it is interesting to note that the proportion of genuine requests steadily increased from program to program in this order: core French (.16), extended French (.37), French immersion (.42), and ESL (.52).

There were important differences between programs in the category of Sustained speech. Teacher turns in core French were rarely sustained; only 28% of core French teacher turns were longer than a sentence. As in the case of Genuine requests, the proportion of sustained teacher speaking turns in classrooms in the other programs increased in the order: core French (.28), extended French (.52), French immersion (.57), and ESL (.61). The difference in proportion of sustained teacher turns between core French and the remaining three programs was significant ($F(3,14) = 5.37; p<.05$).

The final communicative feature of teacher talk, Incorporation of student utterances, reflects the ways in which teachers reacted to student utterances. As indicated in Figure 1, teachers in all programs most frequently used Comments, such as "Good" and "Right," in reacting to students' utterances; Paraphrase was used the least. One interesting difference among programs involved the use of Expansions and Elaborations. These categories occurred extremely rarely in core French and ESL, but they were used at least to some degree (although the differences were not statistically signficant) in extended French and French immersion. Despite the argument that elaborations and expansions contribute to first language development, teachers in this study rarely built on student responses to develop a topic or engaged students in further discourse. It has to be remembered, however, that the sample was extremely small and may not have been representative.

Student verbal interaction. As Figure 2 (p. 53) indicates, student verbal interaction was almost always in the target language. It should be pointed out, however, that students generally used the target language only while their teacher exercised control over classroom activities. During seat work, most interaction occurred in the native language. This is not reflected in the present data, since at those times the tape recorder was usually turned off.

Students in core French gave significantly fewer Unpredictable responses (.14) than did students in the other three programs ($F(3,14) = 4.38; p<.05$). The greater proportion of unpredictable responses in extended French (.41) and French immersion (.49) can be partially attributed to the introduction of subject matter (i.e., history). When the focus is on meaning and on topics other than the language code, the opportunities for teacher questions to which more than one answer is acceptable increase.

Core French also differed from the other programs in terms of length of student speaking turns. The majority (.58) of student speaking turns in core French were Ultraminimal; in the other three programs, student turns were much more often Minimal: .56 in extended French, .46 in French immersion, and .44 in ESL. Sustained turns were almost non-existent (.03) in core French; they increased slightly in extended French (.11) and rose to .29 and .31 in ESL and French immersion, respectively.

The final set of data in Figure 2 reflects the degree of restriction imposed on the linguistic forms which students could use in producing target language utterances. Unrestricted utterances were very infrequent in core French (.07) but increased in the order: ESL (.34), extended French (.47), and French immersion (.71).

The remaining three categories — Reaction to message/code, Incorporation of preceding utterances, and Discourse initiation — occurred extremely rarely in student verbal interaction in all four programs and are therefore not included in Figure 2.

4. DISCUSSION

The results of the validation study showed that many of the descriptive categories introduced in Part I of the COLT were capable of differentiating between the four L2 programs observed. The categories of Content and Materials were particularly revealing in this regard.

Expectations about the distinguishing characteristics of each program were largely supported. The main exception was the ESL program, in which a great deal of group work in which students would discuss topics other than the language code was expected. Instead, a strong focus on grammar and vocabulary was found, even during group work, and students were frequently involved in individual seat work which did not foster communication. When communication did take place during seat work activities, it was generally in the students' first language. One possible reason for the focus on form in the ESL classes is that the ESL learners in this study, unlike the FSL learners, had considerable opportunity for acquisition outside the classroom and that because of this, the ESL teachers may have felt that the language code was the appropriate focus for the classroom.[6]

To characterize each program according to the degree to which it was communicatively oriented — that is, to place each program on a 'communicative continuum' — the investigators decided to select those features which are frequently mentioned in the literature on communicative language teaching and to assign scores from 1 to 5 depending on the percentage of time spent on each. The selected categories were as follows:

- Group work
- Focus on meaning (including management and other topics) and any combinations of form and the other content categories
- Topic control by teacher and students or student alone
- Use of extended text

- Use of semi- and non-pedagogic materials

The scores were based upon an interval scale: 0-19 percent of class time equalled a score of 1, 20 to 39 percent a score of 2, etc. Thus a class which spent 15 percent of class time on group work, 45 percent on other topics, 10 percent on activities controlled by students, 90 percent on extended text, and 15 percent with non-pedagogic text received individual scores of 1+3+1+5+1, yielding a total of 11.

When these calculations are made on the data, the following order was obtained:
Core French (6)
ESL (7)
Extended French (10)
French immersion (12)

In other words, core French was the least 'communicative' in terms of the categories, and immersion the most. ESL and extended French occupied a place in between. It should be emphasized again that the purpose of this study was to determine whether the COLT scheme was capable of capturing differences in the communicative orientation of different types of classroom. The results reported here cannot be interpreted as an evaluation of the L2 programs observed, since the data base was far too small. Moreover, one can assume that there would be considerable variability between teachers within programs - a fact which has not been allowed for in this report.

While results of the Part II analysis confirmed some of the findings of other studies on classroom interaction (e.g., Naiman, Fröhlich, Stern, and Todesco 1978, Sinclair and Coulthard 1975, Wells 1981) — for example, that students usually have the exclusive role of responding to questions, which are generally pseudo-requests, and that students rarely interact with each other in teacher-centered classrooms — there were some interesting differences among the programs observed in this study. In particular, students in immersion classes, where subject-matter instruction in the L2 is part of the curriculum, were given more opportunity for unrestricted language use, for sustained speech, and for giving unpredictable information. In contrast, students in core French classes were required to give predictable responses in restricted form and of ultraminimal length. The extended French and ESL classes tended to be situated in between core and immersion classes. These findings were consistent with the ordering of classes along the communicative continuum obtained in the analysis of the Part I features.

In conclusion, the validation study was conducted to examine whether the COLT observation scheme, which was derived from a model of communicative competence and a review of current issues in communicative language teaching, was capable of capturing differences in the communicative orientation of four second language programs. The results provided preliminary evidence that the scheme is capable of doing so --the programs did indeed differ in their communicative orientation. The development of an observation scheme capable of capturing the characteristics of different types of classroom is an important step towards identifying what makes one set of instructional techniques more effective than another.

Footnotes

1. In this chapter we discuss the categories of the observation scheme as they were established during the development study. A number of modifications were made for the purpose of the core French observation study, and these will be discussed in Chapter 4.

2. A teacher may select a topic and then give the students a great deal of freedom in developing the topic, for example: "Write a short paragraph about your impressions when you first came to Canada". In such cases, Teacher/student control would be checked off.

3. The coders found that Use of materials frequently overlapped with Topic control. Furthermore, it proved difficult to find a satisfactory definition for Use of materials, so in the revised version of COLT this category has been deleted.

4. It should be noted that materials developed for teaching/learning purposes, not for second language learners but for native speakers of the target language, were coded as Pedagogic. This applies particularly to extended French, French immersion, and ESL programs. In the core French observation study, such materials were coded separately.

5. For those communicative features which consist of two categories, only one of the proportions has been graphed.

6. Both in the validation study and in the core French observation study, there was insufficient time to conduct interviews with teachers. In the future, it will be important to make sure that there is time available for teacher interviews.

References

Aschner, M.J., et al. **A System for Classifying Thought Processes in the Context of Classroom Verbal Interaction.** Urbana, Ill., University of Illinois, 1965.

Bellack, A.A., Hyman, R.T., Smith, F.L., Jr. Kliebard, H.M. **The Language of the Classroom.** Final report, USOE Cooperative Research Project, No. 2103. New York: Teachers College, Columbia University, 1966.

Bialystok, E., Fröhlich, M., Howard, J. Studies on second language learning and teaching in classroom settings: strategies, processes and functions. Toronto: Ontario Institute for Studies in Education, unpublished report, 1979.

Breen, M.P. and Candlin, C.N. The essentials of a communicative curriculum in language teaching. **Applied Linguistics,** 1980, **1,** (2), 89-112.

Breen, M.P. Authenticity in the language classroom. **Bulletin of the Canadian Association of Applied Linguistics.** 1982, **4,** (2), 7-24.

Brown, H.D. **Principles of Language Learning and Teaching.** Englewood Cliffs, N.J.: Prentice-Hall, 1980.

Brumfit, C.J. Teaching pupils how to acquire language: some comments on the positions implicit in the preceding two articles. **ELT Documents,** 1976, **3,** 24-27.

Brumfit, C.J. Teaching the 'general' student. In Johnson, K. and Morrow, K. (eds.) **Communication in the Classroom.** London: Longman, 1981.

Byrne, Don. **Functional Comprehension.** London: Longman, 1977.

Canale, M. From communicative competence to communicative language pedagogy. In Richards, J. and Schmidt, R. (eds.). **Language and Communication.** London: Longman, 1983.

Canale, M. and Swain, M. Theoretical bases of communicative approaches to second language teaching and testing. **Applied Linguistics,** 1980, **1,** (1), 1-47.

Corder, 1971. Idiosyncratic dialects and error analysis. **IRAL,** 1971, **9,** 147-160.

Cross, T.G. Mother's speech and its association with rate of linguistic development in young children. In Waterson, N. and Snow, C.E. (eds.) **The Development of Communication.** Chichester: Wiley, 1978.

Davis, O.L., Jr. and Tisley, D.C. Cognitive objectives revealed by classroom questions asked by social studies student teachers. In Hyman, R. T. (ed.) **Teaching: Vantage Points for Study.** Philadelphia: J.B. Lippincott, 1968.

De Villiers, P.A. and de Villiers, J.G. **Early Stages.** Cambridge: Harvard University Press, 1979.

Dunkin, M.J. and Biddle, B.J. **The Study of Teaching.** New York: Holt, Rinehart and Winston, Inc., 1974.

Ellis, R. and Wells, C.G. Enabling factors in adult-child discourse. **First Language,** 1980, 1, 46-82.

Fanselow, J.F. Beyond Rashomon - conceptualizing and describing the teaching act. **TESOL Quarterly,** 1977, 11 (1), 17-39.

Flanders, N.A. **Analyzing Teaching Behavior.** Reading, Mass.: Addison-Wesley, 1970.

Fletcher, Mark, and Hargreaves, Roger. **Evans Functional Units: Activating Vocabulary.** London: Evans Brothers Ltd., 1980.

Forsyth, I.J. Patterns in the discourse of teachers and pupils **CILT Reports and Papers,** 1974, **10, 77-99.**

Hatch, E. Discourse analysis and second language acquisition. In E. Hatch (ed.) **Second Language Acquisition.** Rowley, Mass.: Newbury House, 1978.

Hymes, D. 1972 On communicative competence. In Pride, J.B. and Holmes, J. (eds.) **Sociolinguistics.** Harmondsworth, England: Penguin Books, 1972.

Johnson, K. **Communicative Syllabus Design and Methodology.** Oxford: Pergamon Press, 1982.

Johnson, Keith, and Morrow, Keith. **Approaches.** Cambridge: Cambridge University Press, 1979.

Johnson, Keith, and Morrow, Keith, (eds.) **Communication in the Classroom.** London: Longman, 1981.

Long, M.H., Leslie, A., McLean, M. Castanos, F. Doing Things with words - verbal interaction in lockstep and small group classroom situations. In Fanselow, J.F. and Grimes, R.H. (eds.) **On TESOL '76.** Washington: TESOL, 1976.

Long, M.H. Inside the "black box": methodological issues in classroom research on language learning. **Language Learning,** 1980, **30,** 1-42.

Long, M.H. Training the second language teacher as classroom research. Paper presented at the 34th Annual Round Table on Languages and Linguistics. Georgetown University, Washington, D.C., March 1983.

MacLure, M. and French, P. A comparison of talk at home and school. In Wells, G. (ed.) **Learning Through Interaction.** Cambridge: Cambridge University Press, 1981.

Macnamara, J. Nurseries, streets and classrooms: some comparisons and deductions. **Modern Language Journal,** 1973, 57, 250-254.

McEwen, N.Z. An exploratory study of the multi-dimensional nature of teacher-student verbal interaction in second language classrooms. Unpublished doctoral dissertation, University of Alberta, 1976.

McNeill, D. Developmental psycholinguistics. In Smith, F. and Miller, G.A. (eds.) **The Genesis of Language: A Psycholinguistic Approach.** Cambridge, Mass.: MIT Press, 1966.

Mehan, H. "What time is it, Denise?": asking known information questions in classroom discourse. **Theory into Practice**, 1979, 18, (4), 285-294.

Mitchell, R., Parkinson, B., Johnstone, R. **The Foreign Language Classroom: An Observational Study.** Stirling Educational Monographs No. 9, Dept. of Eduction, University of Stirling, Scotland, 1981.

Morrow, K. **Technique of Evaluation for a Notional Syllabus.** London: The Royal Society of Arts, 1977.

Morrow, K. Principles of communicative methodology. In Johnson, K. and Morrow, K. (eds.) **Communication in the Classroom.** London: Logman, 1981.

Moskowitz, G. **The Foreign Language Teacher Interacts.** Chicago: Association for Productive Teaching, 1970.

Moskowitz, G. Interaction analysis - a new modern language for supervisors. **Foreign Language Annals**, 1971, 5 (2), 211-221.

Munby, J. **Communicative Syllabus Design.** Cambridge: Cambridge University Press, 1978.

Naiman, N., Fröhlich, M., Stern, H.H., Todesco, A. **The Good Language Learner.** Toronto: Ontario Institute for Studies in Education, 1978.

Phillips, M. and Shettlesworth, C. Questions in the design and implementation of courses in English for specialised purposes. Paper presented at the 4th AILA Conference, Stuttgart, 1975.

Richards, J.C. A noncontrastive approach to error analysis. In Oller, J. and Richards, J. (eds.) **Focus on the Learner: Pragmatic Perspectives for the Language Teacher.** Rowley, Mass.: Newbury House, 1973.

Scherer, George A.C. and Wertheimer, M. **A Psychological Experiment in Foreign Language Teaching.** New York: McGraw-Hill, 1964.

Simon, A. and Boyer, E.G. **Mirrors for Behaviour III: An Anthology of Classroom Observation Instruments.** Wyncote, Penn.: Communication Materials Center, 1974.

Sinclair, J.M. and Coulthard, R.M. **Towards an Analysis of Discourse.** London: Oxford University Press, 1975.

Smith, P.D. **A Comparison of Cognitive and Audiolingual Approaches to Foreign Language Instruction: The Pennsylvania Foreign Language Project.** Philadelphia: The Centre for Curriculum Development, 1970.

Snow, C. and Ferguson, C.A. (eds.) **Talking to Children: Language Input and Acquistion.** Cambridge, England: Cambridge University Press, 1977.

Stern, H.H. The formal-functional distinction in language pedagogy: a conceptual clarification. In Savard, J.G. and Laforge, L. (eds.) **Proceedings of the 5th International Congress of Applied Linguistics, Montreal, 1978.** Publication A16 of the International Centre for Research on Bilingualism. Quebec: Les Presses de l'Universite Laval, 1981, 425-455.

Wells, C.G., Montgomery, M., MacLure, M. Adult-child discourse: outline of a model of analysis. **Journal of Pragmatics**, 1979, 3, 337-80.

Wells, G. **Learning through Interaction. The Study of Language Development.** Cambridge: Cambridge University Press, 1981.

Widdowson, H.G. **Teaching Language as Communication.** Oxford: Oxford University Press, 1978.

Wilkins, D.A. **Notional Syllabuses.** London: Oxford University Press, 1976.

Withall, J. The development of a technique for the measurement of social-emotional climate in classrooms. **Journal of Experimental Education**, 1949, 17, 347-361.

Ullmann, R. and Geva, E. The target language observation scheme (TALOS). York Region Board of Education Core French Evaluation Project, Ontario Institute for Studies in Education, 1982 (mimeo).

TABLE 1

Participant Organization: Mean Percentages of Observed Time by Program

	Whole Class					
	T ↔ S/C	S ↔ S/C	Choral	Group	Individual	Group/ Individual
Core French (4)	58.49	2.72	14.40	5.01	19.38	0
Extended French (2)	70.48	17.20	0	0	12.32	0
French Immersion (2)	60.90	17.32	2.73	0	19.05	0
ESL (5)	21.28	11.05	1.28	10.00	43.02	13.37

TABLE 2

Content: Mean Percentages of Observed Time by Program

		Language		Other Topics				Combinations			
	Management	Form	Socio.	Narrow	Limited	Broad	Total	Form/ Socio.	Form/ Limited	Form/ Broad	Socio./ Limited
Core French (4)	2.37	58.44	0	1.67	22.84	3.38	27.89	5.50	5.80	0	0
Extended French (2)	9.45	25.10	0	0	9.55	31.00	40.55	0	0	24.90	0
French Immersion (2)	4.75	14.35	11.12	0	11.20	51.33	62.53	0	0	7.25	0
ESL (4)	5.85	66.43	0	1.18	7.83	7.51	16.52	0	3.37	7.50	.33

TABLE 3

Topic Control: Mean Percentages of Observed Time by Program

	Teacher	Teacher/Student	Student
Core French (4)	93.89	6.11	0
Extended French (2)	88.72	11.28	0
French Immersion (2)	80.02	19.33	.65
ESL (4)	91.08	8.92	0

TABLE 4

Student Modality:
Mean Percentages of Observed Time by Program

	Modalities*					Combinations						
	L	S	R	W	Other	LS	LSR	LR	RW	LSRW	LW	SR
Core French (4)	7.46	1.09	0	1.66	2.68	38.30	24.78	3.58	16.96	2.13	1.36	0
Extended French (2)	13.70	0	1.06	0	2.40	19.52	44.40	4.58	11.23	1.75	0	1.34
French Immersion (2)	12.87	0	1.37	2.25	0	32.50	29.57	4.17	6.77	10.50	0	0
ESL (4)	2.85	0	.84	3.52	1.68	24.33	24.45	3.68	35.31	2.64	.70	0

* L = Listening; S = Speaking; R = Reading; W = Writing

TABLE 5

Type of Materials:
Mean Percentages of Observed Time by Program

	Text				Combinations			
	Minimal	Extended	Audio	Visual	Min. Text/ Visual	Ext. Text/ Audio	Ext. Text/ Visual	No Materials Used
Core French (4)	43.06	11.31	.68	18.23	15.93	1.01	0	9.76
Extended French (2)	35.11	38.19	3.75	5.28	0	0	1.35	16.32
French Immersion (2)	31.20	50.90	0	4.10	0	0	7.73	6.07
ESL (4)	52.29	34.73	0	1.08	0	0	0	11.90

TABLE 6

Source of Materials:
Mean Percentages of Observed Time by Program[1]

	Pedagogic[2]	Semi-Pedagogic	Non-Pedagogic
Core French (4)	83.69	4.90	1.69
Extended French (2)	72.88	5.53	5.30
French Immersion (2)	67.56	2.25	24.13
ESL (4)[3]	63.99	2.84	15.75

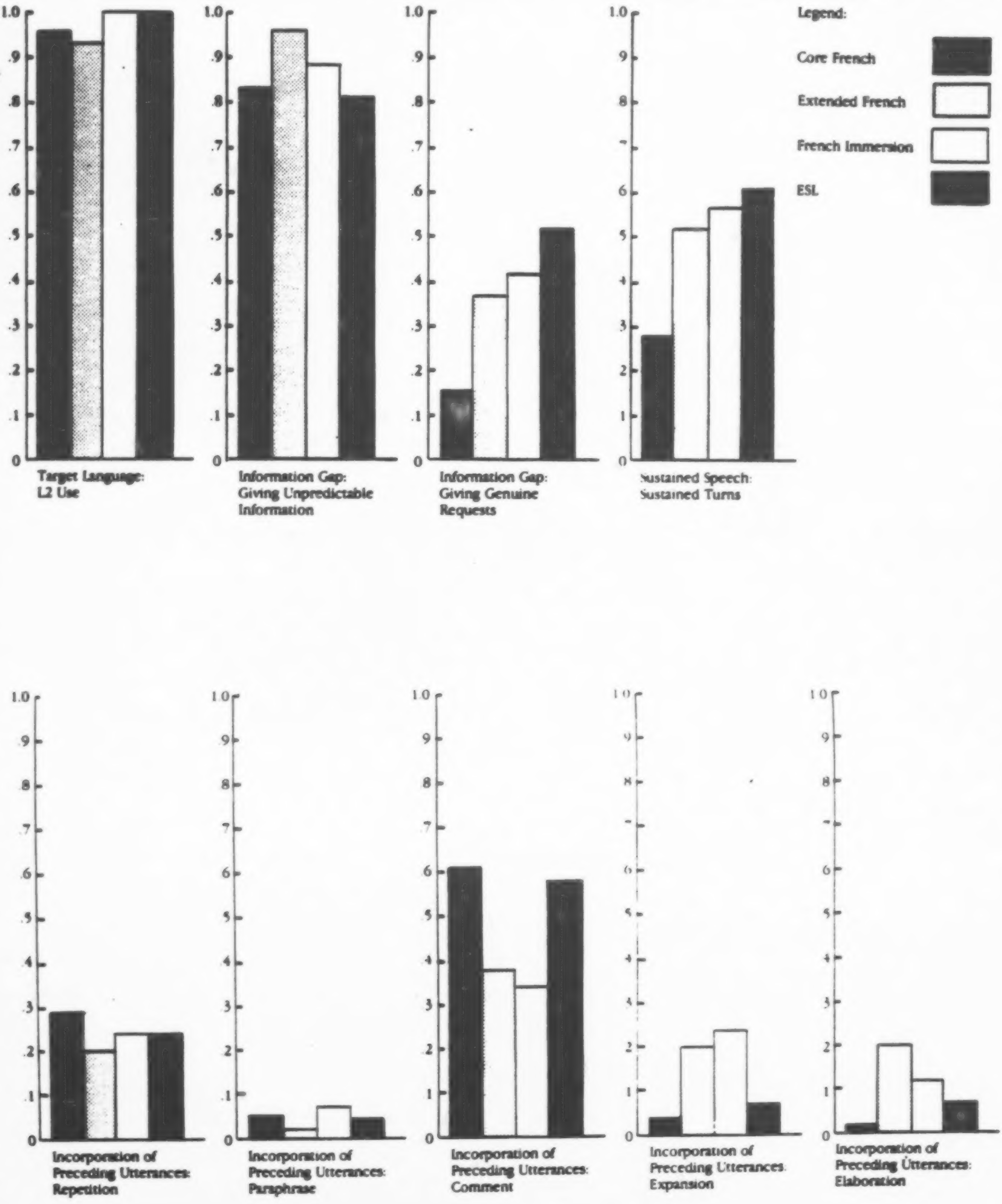
[1] Percentages, calculated from total class time observed, do not add up to 100 percent because materials were not used all the time.

[2] These figures also include materials developed for native speakers of the target language; this applies to the extended French, French immersion, and ESL programs.

[3] Data for 5.52 percent of the time observed are missing.

FIGURE 1

Category as Proportion of Communicative Feature for Teacher Talk, by Program



Target Language: L2 Use

Information Gap: Giving Unpredictable Information

Information Gap: Giving Genuine Requests

Sustained Speech: Sustained Turns

Incorporation of Preceding Utterances: Repetition

Incorporation of Preceding Utterances: Paraphrase

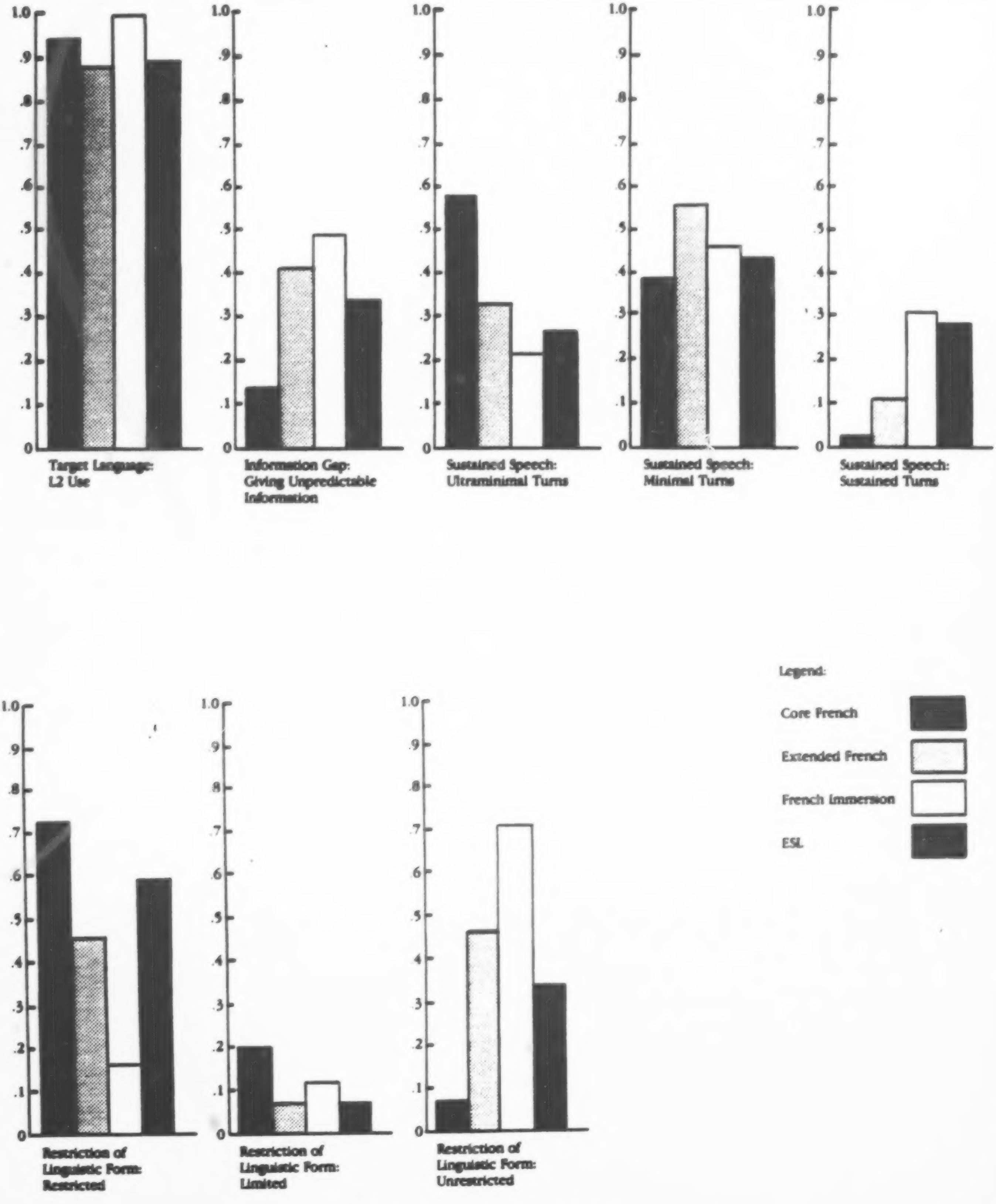
Incorporation of Preceding Utterances: Comment

Incorporation of Preceding Utterances: Expansion

Incorporation of Preceding Utterances: Elaboration

FIGURE 2

Category as Proportion of Communicative Feature for Student Talk, by Program



APPENDIX A
COLT Observation Scheme (Part A)

SCHOOL ______________ **GRADE(S)** ______________ **DATE** ______________

TEACHER ______________ **LESSON (Minutes)** ______________ **OBSERVER** ______________

SUBJECT ______________

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
TIME	ACTIVITIES	PARTIC. ORGANIZATION							CONTENT																								STUDENT MODALITY					MATERIALS									
		Class			Group		Comb.		Man.		Language					Other Topics														Topic Control								Type							Use		
																Narrow				Limited					Broad													Text									
		T – S/C	S – S/C	Choral	Same	Different	Individual	Gr/Ind.	Procedure	Discipline	Implicit	Form	Function	Discourse	Soc. Ling.	Classroom	Stereo.(s)	Pers./Bio.	Other	Personal	Rout./Soc.	Fam./Com.	School T.	Other	Abstract	Pers./Ref.	Imagination	World T.	Other	Teacher	Teacher/Stud.	Student	Listening	Speaking	Reading	Writing	Other	minimal	extended	Audio	Visual	Pedagogic	Semi-Pedag.	Non-Pedag.	High Control	Semi Control	Min. Control

© J. P. B. Allen, M. Frohlich and N. Spada, OISE 1983

189

COLT Observation scheme (Part B)

TEACHER VERBAL INTERACTION																
Communic. Features		Target Lang.		Information Gap				Sust. Speech		Reaction Co/Mes	Incorporation of S. Utterances					
				Giving Info.		Request Info.										
No talk	Off talk	L_1	L_2	Predict.	Unpred.	Pseudo	Genuine	Minimal	Sustained	Explicit Code Reaction	No Incorp.	Repetition	Paraphrase	Comment	Expansion	Elaboration

STUDENT VERBAL INTERACTION																				
	Target Lang.			Information Gap				Sust. Speech			Form Restr.			Reaction Co/Mes.	Incorporation of S/T Utterances					
				Giving Info.		Request Info.														
Choral	L_1	L_2	Disc.-Initiation	Pred.	Unpred.	Pseudo	Genuine	Ultraminimal	Minimal	Sustained	Restricted	Limited	Unrestricted	Explicit Code Reaction	No Incorp.	Repetition	Paraphrase	Comment	Expansion	Elaboration

© J. P. B. Allen, M. Fröhlich and N. Spada, OISE 1983

Chapter 4

THE CORE FRENCH OBSERVATION STUDY

Patrick Allen, Susanne Carroll, Jud Burtis and Vince Gaudino

1. INTRODUCTION

The results of the validation study provided evidence that the COLT observation scheme is capable of capturing differences in the pedagogic orientation of different second language programs. Crucial questions, however, remain unanswered: Will the COLT prove equally useful in capturing differences within a single program incorporating a fairly homogeneous sample? Will all activities prove to be equally important in defining the nature of classroom treatment? Are all activities equally relevant to determining learning outcomes, or are certain activities or certain combinations of activities more likely to have an impact on specific aspects of proficiency? Thus, will students from classes with a relatively strong analytic focus score higher on measures of grammatical competence than students from classes with a relatively strong experiential focus?[1] Are students from experiential classes more likely to score higher on measures of discourse and sociolinguistic competence because they have had more classroom experience using language for communication and producing extended text? Only a study which compares instructional differences across classes within the same program and relates these to differences in proficiency can hope to provide an answer to such questions.

The aim of the process-product study reported here was to relate instructional differences in the core French program at the grade 11 level to differences in the communicative competence of the students. In other words, we wanted to investigate how instructional differences affect learning by relating aspects of the classroom environment to proficiency measures. We began our study with certain minimal assumptions:

- student and teacher behaviour would vary from one class to another within the sample of core French classes;
- these differences in classroom behaviour would be characterizable in terms of the COLT categories and would be significant enough in terms of both the types of activities occurring and the time devoted to each to permit a ranking of classes along an experiential-analytic scale;
- significant differences in classroom behaviour would correlate with specific aspects of second language proficiency, namely grammatical, discourse, sociolinguistic and strategic proficiency.

2. THE SAMPLE

2:1 Selection of the Classes

The sample consisted of eight grade 11 core French classes which were selected in a two-part process. Initially, the personnel of three Metropolitan Toronto school boards were asked to suggest the names of a number of teachers using either experiential or analytic approaches in their classes and who might be prepared to participate in the study. As a result, a preliminary list of 13 teachers was compiled. Subsequently, in the spring of 1984, all the teachers were observed by one of the team's researchers who had been instrumental in the design of the COLT. A final selection of eight classes was made on the basis of these informal observations. No analysis of the classes using the COLT was done at this time, but we were working on the assumption that we would eventually be able to divide the classes into two distinct groups, namely (a) an experiential or functionally-oriented group, and (b) an analytic or structurally-oriented group. The final classification and ranking of the classes depended on the outcome of the observations conducted throughout the course of the study using the COLT, leaving open the possibility that classes might be more-or-less experiential, or more-or-less analytic according to a theoretically-defined absolute scale.[2]

2:2 Characteristics

The core French program was chosen as the context for the study because it consists of a relatively homogeneous group of students with respect to the amount of previous instruction time and exposure to French outside of the class. In Ontario, French is now a compulsory subject up to Grade 9 and studies usually begin in Grade 4. The grade 11 students who were our subjects were studying French as an optional subject and they had had approximately eight years of prior core French instruction. Furthermore, since participating classes were drawn from three school boards in Metropolitan Toronto, all students lived in an environment where English was the dominant language and where opportunities to use French for communicative purposes outside the classroom were severely limited. Thus, the expectation was that the students' knowledge of French and communicative skills would derive largely from the school environment.

Three classes came from each of two boards and two classes came from the third (see Table 1, p. 98). Three classes received forty minutes of instruction five times a week while five classes received seventy minutes three times a week. On a weekly basis, the grade 11 core French classes received an average of two hundred minutes of instruction.

The average class size was 23.5 students, including one small class of ten students (see Table 2, p. 99). The average student age was 16.9 years. The classes consisted predominantly of monolingual anglophones (see Table 3, p. 100). Class 6 was an exception to this generalization in that only 15% of the students were monolingual anglophones. Students came from a mixture of socio-economic backgrounds.

Information about students' contact with French outside the classroom was obtained via a questionnaire which was given twice, once in October at the time of pretesting, and again in April during post-testing (for copies of the questionnaires see Appendix A, pp. 127-136, and for a summary of the results see Table 4, p. 101). About 7% of students in the sample said that they had previously been enrolled in an immersion program. Only about one-third of these stayed in the program longer than one year. About 4% of the sample reported that they had previously been in an extended French program. Most of these stayed in the program from two to four years.

In the questionnaire the students were also asked to rate themselves on a 5-point scale (1 = not at all, 5 = extremely well) as to how well they understood, spoke, read and wrote French. None of the students reported that they understood not at all, most (49%) reported they understood fairly well, and 24% reported they understood quite well. With regard to speaking, 38% said they spoke with some difficulty, 49% reported they spoke fairly well and 8% quite well. In reading, most students (41%) reported that they read fairly well, 24% said they read with some difficulty and 29% said they read quite well. Writing was found to be generally more difficult. About 2% of the students said they wrote not at all, 40% said they wrote with some difficulty, 42% said fairly well, and 15% quite well. No one claimed they wrote extremely well in French.

Students were also asked about their contact with French during the current year outside of French classes at school. Students were asked to indicate on a 5-point scale how often (1 = never, 5 = daily) they participated in twelve activities related to French. For most of the activities some 60% of students reported they never or hardly ever participated, and on average only about 5% said they participated frequently. The one exception was reading the French labels and advertisements on packages. Most students (74%) reported they read these sometimes or frequently, and only 22% said they never or hardly ever read them.

Only one student reported taking French classes in addition to the classes at school (less than 1% of the sample). Only one student reported participating in any extra-curricular French activities during the current year. Most students (55%) reported that they never or hardly ever had an opportunity to use French outside of school, and 37% said they sometimes had an opportunity. Only 8% said they used French often or extremely often outside the school context.

Some 37% of students reported that they had spent no time in a French-speaking area during the past five years, 35% reported spending up to two weeks, 15% two to six weeks, and 13% more than six weeks.

About 22% of students reported that they had participated in some form of extra-curricular French activity during the previous five years. For most students the total contact time was two weeks or less. About a quarter of the students reported contact of up to four weeks, and most of the remaining students reported contact of more than six weeks.

In order to determine whether contact with French outside the classroom had any effect on the achievement scores of the analytic and experiential groups, these groups were subcategorized into students having high contact, and students having low contact with French. Students were considered to have high contact if they had previously participated in an immersion or extended French program, if they had spent more than six weeks in a French-speaking community during the past five years, or if they spoke French with native speakers frequently or daily during the current year. A total of 55 students were found to have had high contact with French outside the classroom.

A two-way analysis of co-variance was calculated by group (analytic or experiential) and contact (high or low), using the pre-test score as the covariant, for post-test scores on the four written measures and the four oral measures used in the study. The analysis revealed that for this population of students high contact with French outside the classroom had virtually no effect on achievement scores.[3]

The picture that emerges, then, is of a relatively homogeneous group of L2 learners at the grade 11 level. The students had had approximately eight years of instruction, mainly in the core French program, and in most cases they had little opportunity to use the target language outside their regular core French classes.

3. GENERAL PROCEDURES

3:1 Pre-tests

All classes were given, over two sessions, a series of pre-tests which consisted of (1) a multiple choice grammar test (allotted time: 20 minutes), (2) a written exercise taking the form of a note whose function was to elicit an informal request (henceforth 'the note' -- allotted time: 15 minutes), (3) a second written exercise consisting of a formal request to be expressed in the shape of a letter (henceforth 'the letter' -- allotted time: 15 minutes), (4) a listening comprehension test with multiple choice answers (henceforth the LCT -- allotted time: 20 minutes). Copies of all the written tests appear in Appendix B, pp. 137-160. During one of the testing sessions, students also completed the contact questionnaire. The students were told at each testing session by tester 1 (a native speaker of French) that anonymity would be ensured and that the results of the tests would have no effect on their school marks. Instructions were given both in French and in English.

At the same time that the whole-class testing was being conducted, a subset of students, whose names had been randomly selected from the class list, were interviewed orally (by tester 2, a fluently bilingual anglophone) in a separate room.[4] Once again, students were told in French that anonymity would be guaranteed, that the interview would not affect their marks and that they had been selected randomly. Although the entire interview was conducted in French, students were aware that the interviewer spoke English as well.[5] In all, 48 students (six from each class) were pre-tested during structured interviews which lasted anywhere from ten to thirty minutes. In general, the more difficulty a student had in either understanding the interview questions, or in responding to them, the longer the interview lasted. Interview length, however, is not a reliable indicator of ability since some students responded with *I don't know* or *Je sais pas*, forcing the interviewer to go quickly on to the next question, while other students, who were quite comfortable speaking, chatted on about a topic for several minutes. Although some shyness and hesitation were to be expected, overall the interviews did not seem to bother the students. One or two students expressed a certain nervousness about the 'test' nature of the interview.[6] All students appeared to do their best to reply and some were happy simply not to be in class. One or two individuals expressed pleasure at having the chance to try out some 'real French'. Briefly, although the presence of the testers was clearly a deviation from normal routine, the students accepted it and made them welcome.

All students were interviewed using the same schedule (see Appendix C, pp. 161-163) and the interviews were recorded on audiocassettes. Students spoke into a clip microphone so that the quality of sound on the interviews was excellent. The interviews were then manually transcribed by the interviewer (a trained linguist) in an augmented 'normal' orthography.[7] The transcriptions were revised twice -- once by a fluent speaker of French and English whose mother tongue was Portuguese, once by the original transcriber/interviewer. The transcriptions were then entered onto computer and partially coded at the same time.[8] The coding and data entry were done by the interviewer.

3:2 Observations

Subsequent to the pre-testing sessions, and before the post-testing, each class was visited four times for an observation. These observations occurred in October, January, March and April, i.e. as far as possible, they were organized to extend over a full school year. The observation periods lasted the full time of the class and so lasted either forty or seventy minutes. They were also recorded on audiotape, using two microphones which were placed in such a way as to capture both the teacher's speech and the students' speech. The observer sat at the back of the class and tried to be as inconspicuous as possible. The tapes were transcribed after the observations in ordinary French orthography by the principal observer (tester 1). Out of 32 observations (eight classes x four observations), 20 were carried out by the principal observer who had been previously trained on the validation study. Four were conducted by tester 2. On eight occasions, the two observers observed together and discussed the observations and coding immediately following the class.

The coding procedure was similar to the one used in the development and validation studies. Part I of the COLT scheme, describing instruction at the level of activity, was used during the class time and was filled out in the class by the observer. Part II coding was done after the observation, using the transcriptions and/or the audiorecording of the observed class. A time-sampling procedure within each activity identified in Part I was used. Coding began at the beginning of each activity, lasted for one minute, and resumed after a two-minute interval. During the one-minute coding periods, the frequency of occurrence of each category of the communicative features of teacher and student interaction was recorded. A number of issues arising out of the Part II coding were discussed with the researcher who had conducted the observations during the validation study and who had trained our principal observer-tester.[9]

A supplement to the observation scheme, referred to as 'COLT Part III', was used to obtain more detailed information about the nature and organization of form-oriented activities in the classroom. COLT Part III consisted of nineteen yes/no questions related to how structure was taught (namely, in terms of oral production, reading, listening or writing -- see Appendix D, pp. 164-166). It was completed immediately after each classroom observation. It was discovered that COLT Part III did not add substantially to the information provided by Parts I and II of the COLT and the teacher questionnaire. These data were omitted from the final analysis.

3:3 Post-tests

In May, the classes were given the same written tests, under the same circumstances with the same two testers. Instructions remained the same. The same students who were orally pre-tested were selected for oral post-testing. Due to absenteeism, however, the number of students interviewed fell from 48 to 43. The questions on the schedule were modified slightly to make them more topical; the linguistic focus of each question nevertheless remained the same from pre-test to post-test. Questions from both interview schedules are incorporated into Appendix C.

3:4 Teacher Questionnaire

A teacher questionnaire was devised to obtain information about teaching/learning activities throughout the year, including those occasions when the observers were not present (a copy appears in Appendix E, pp. 167-180). Information was elicited about the use of texts and supplementary materials, the organization of writing, reading and listening/speaking activities, the use of activities with an explicit focus on grammar, discourse and sociolinguistics, methods of correcting student errors, homework assignments given during the year, and use of L2 before and after the actual French period. The questionnaire was handed out to teachers on the next to last observation session, and was collected at the final session. By this means, a 100% rate of return was obtained.

4. THE COLT OBSERVATION SCHEME: PROCEDURES

4:1 Modifications to the Observation Scheme

For the purpose of the process-product study, the following modifications were made to the COLT observation scheme described in Chapter 3:

Individual work. This feature was modified to include two subcategories: Same (all students working on the same task) and Different (students working on different tasks).

Topic control. This category was modified to allow for the possibility that topic selection by the teacher may be done in conjunction with a textbook.

Source/purpose of materials. The labels Pedagogic, Non-pedagogic and Semi-pedagogic were replaced by L2 (materials specifically designed for FSL teaching), L1 (materials originally intended for francophone L1 or non-school purposes), and L1-adapted (utilising L1 materials or real-life objects and texts, but in a modified form). The subcategory Student made was added to provide information about materials which were produced by the students themselves.

Use of materials. As a result of further experience in using the observation scheme, it became apparent that Use of materials in Part I provided essentially the same information as Relative restriction of linguistic form in Part II. In the process-product study, therefore, the category Use of materials was deleted.

Requesting information. The labels Pseudo request and Genuine request were changed to Display request and Information request respectively.

Reaction to code or message. This feature was modified to include two subcategories: Explicit reaction to code, and Explicit reaction to message.

Incorporation of preceding utterances. The subcategory No incorporation was deleted and two new categories added: Correction (i.e. correction of previous utterance/s) and Clarification (request for clarification of preceding utterances/s).

Relative restriction of linguistic form. The subcategory Limited restriction was deleted.

A copy of the observation schedule as used in the process-product study, together with a list of revised definitions, will be found in Appendix F, pp. 181-189.

4:2 Coding Procedures

In order to calculate the percentage of observed time for each category under the five main headings of COLT Part I we proceeded as follows. The first parameter in Part I was open-ended, i.e., no predetermined descriptors had to be checked off by the observer. Each activity and its constituent episodes were separately described, e.g.: drill, translation, discussion, game (separate activities); teacher introduces dialogue, teacher reads dialogue aloud, students repeat dialogue parts after teacher (three episodes of one activity). During observation, the coder had indicated the beginning and ending times for each activity and episode. The first step was to calculate the time in minutes for each episode (i.e., the smallest unit observed). For each episode, the coder indicated which features of classroom interaction had been observed under each major heading. A time value in minutes was then assigned to each Part I category. If two or more categories were marked under one major heading, the coder indicated which category constituted the major focus. In such cases, the primary category received credit for the entire length of time the episode lasted. If two or more categories were considered to be of equal importance, the time was divided equally among them. This procedure was followed for all four observations.

The next step was to sum the various time values assigned to each Part I category, in order to arrive at the percentage of observed time for each category in each classroom across four observations. The percentage of observed time for each category was then calculated, following a similar procedure. The sum of the categories under Materials did not add up to 100% since materials were not always used during the whole of observed class time. For the other major headings (Participant organization, Content, Content control, and Student modality) the sum of the percentages in each category totalled 100% of observed time.

The coding for COLT Part II was based on an audiotape recording of the class, a time-sampling procedure being followed. As previously described, coding started at the beginning of each observation, lasted for one minute, and was resumed after a two-minute interval. Thus approximately one-third of the observed time for each class was coded under COLT Part II. Each speech turn by teacher or student was coded by placing check marks in the appropriate columns on the coding form. In order to calculate time percentages for Part II it was assumed that all the turns within a given minute of coding were of equal duration. Thus, if there were ten turns coded in a particular minute, then each turn in that minute was deemed to have been one-tenth of a minute long. After the times had been calculated for each turn, appropriate time values were assigned to the columns on the coding sheet. Finally, the percentage of coded time for each Part II category was calculated by dividing the total time coded under each heading by the total number of minutes coded for each school.

When Part II categories were calculated as a percentage of the total amount of class time it was found that the amount of time coded under a particular heading was often rather small. We decided, therefore, that we would also present the data as a proportion of each superordinate category in COLT Part II. For example, if we take the superordinate category Teacher reaction to code or message, and if we find that in a particular classroom the teacher spent 2% of total time reacting to code and 3% reacting to message, then the proportion of time would be .40 for code reaction and .60 for message reaction. In other words, the proportions were calculated by dividing the percentage of time spent under each category by the total percentage of time coded for the superordinate category; in the case of our hypothetical example, 2 ÷ 5 = .40.

As indicated in Chapter 3, the instructional variables selected for examination in the COLT scheme were motivated by a desire to describe as precisely as possible some of the features of interaction which occur in second language classrooms. Our concept of 'pedagogic feature' was derived from current theories of communicative competence, from the literature on communicative language teaching, and from a review of recent research into first and second language acquisition.[10] Wherever possible, the COLT categories were grouped in such a way that each experiential feature was matched by a corresponding analytic feature. The result of this grouping was as follows:

Experiential feature	Analytic feature
COLT Part I	
group activity	whole-class activity
classroom management	
function/discourse/sociolinguistic focus	form focus
broad/limited range of reference	narrow range of reference
student or shared control	teacher control
extended text	minimal text
L1/L1 adapted/student-made materials	L2 materials
COLT Part II	
use of French	use of English
giving unpredictable information	giving predictable information
information request	display request
sustained speech	minimal speech
reaction to message	reaction to code
comment, expansion, clarification, elaboration	correction, repetition, paraphrase
initiation by student	
unrestricted form	restricted form

All Part II categories were coded for student and teacher speech, apart from Initiation by student and Form restriction, which were coded for student speech only. Since the target language was generally used for Classroom management during the observation periods, this category was counted as an experiential feature. The categories Individual seat work, Audio/visual materials, and Student modality were omitted, since it was not possible to determine whether they referred to experiential or analytic activities. In other words, these activities did not in themselves distinguish along the appropriate lines, a given feature being experiential or analytic depending upon other factors.

5: CLASSROOM PRACTICE: FINDINGS

5:1 Ranking of Classes

The eight classes in the sample were ranked on an experiential-analytic scale on the basis of the experiential features listed above (section 4:2). In order to arrive at a score which would permit ranking we took the total percentage of time spent on each of the experiential features in COLT Parts I and II and added the figures together. These calculations yielded the ranking and scores shown in Table 5 (see p. 103).

In order to maximize the differences between experiential classes and analytic classes, the schools were divided into two groups. Several methods of grouping were tried, and all yielded similar results. First, the percentage scores in Table 5 were divided into two groups using the mean as the dividing point. This gave two schools in the experiential group and six schools in the analytic group, rather than two groups of four schools each.

In order to confirm the validity of the grouping and ranking, the above procedure was repeated, using proportion of time spent on experiential features as the basis for the ranking score. This yielded a very similar result. Schools 5 and 2 were still at the top, the mean still fell between school 2 and school 3, and school 4 was still the lowest on the scale. The only difference was that schools 1 and 8 had changed positions on the scale (see Table 6, p. 103).

As a further check the ranking was done a third time, using number of speech turns by teacher or student as the basis for the score. This yielded identical rankings to those obtained using proportion of time, and the mean again fell between school 2 and school 3. Having confirmed the groupings and rankings, we decided that in all future analyses schools 5 and 2 would be regarded as the experiential group (Type E), and the remaining six schools would comprise the analytic group (Type A).

5:2 Characteristics of Type A and Type E Classrooms

As described in the previous section, the eight classrooms in the sample were divided into those which were more analytic than other classes in terms of the total percentage of time spent on analytic activities (Type A), and those which were more experiential than other classes in terms of the total percentage of time spent on experiential activities (Type E). In this section we will provide more details of the differences between Type A and Type E classes, as revealed by Parts I and II of the COLT observation scheme.

Colt Part I

The mean percentages of total observed time for Part I categories are presented in Tables 7A-F (pp. 104-105). The tables show that there are tendencies for Type A and Type E classes to differ in terms of the relative amount of time spent on various types of activity. In the case of content control and type/source of materials these tendencies reach significance, as indicated below.

With regard to content control, Type A students spent significantly more time than Type E classes on activities controlled by the teacher (91.33% and 56.55% respectively, $p < .05$). Type E classes spent 20.21% of their time on activities controlled by the students, while this type of activity did not occur at all in Type A classrooms during the periods of observation. This difference was significant ($p < .001$).

With regard to type/source of materials, Type A classes spent significantly more time on activities involving the use of minimal written text than Type E classes (49.15% and 13.97% respectively, $p < .01$), while Type E classes spent significantly more time on activities involving the use of extended written text (67.52% and 26.91% respectively, $p. < .01$). Furthermore, Type E classes spent 31.18% of their time on student-made materials, while Type A classes spent 5.67%. This difference was significant ($p < .01$).

No statistically significant differences were found in participant organization. However, contrary to expectations, there was a tendency for Type E classes to spend more time than Type A classes on activities involving whole-class interaction (56.55% and 48.19% respectively). No choral work was observed in Type E classrooms, and only a small amount (0.92%) in Type A classrooms. No group work occurred in Type E classrooms during the periods of observation. However, small amounts of group work with groups working on the same task (0.99%) and groups working on different tasks (3.22%) occurred in Type A classrooms.

No statistically significant differences were found in content, student modality, or source of materials. Type A classes spent 56.95% of their time on activities which focused exclusively or primarily on form, while type E classes spent 47.41%. Type E classes, on the other hand, spent more time than Type A classes on activities involving a broad range of reference (29.03% and 15.79% respectively). Explicit focus on function, discourse and sociolinguistics occurred either not at all, or very rarely, during observed time in both types of classroom. The lack of explicit focus in these areas is consistent with results obtained for core French, ESL and French immersion in the validation study.

Type A classes spent more time than Type E classes on materials which were specifically designed to be used in a second language classroom (54.29% and 38.70% respectively). L1 and L1-adapted materials were used infrequently in Type A classes, and not at all in Type E classes during the periods of observation. However, these differences were not significant.

Although, by definition, the differences between Group A and Group E were in the 'right' direction, with Type A classes spending more time on analytic activities and Type E spending more time on experiential activities, it is interesting to note that significant differences were found in only three Part I features (topic selection by teacher/student, production of minimal/extended written text, and use of student-made materials). No significant differences were found in participant organization, explicit focus on language or other topics, student modality, or use of materials specifically designed for the L2 classroom, as opposed to those originally intended for francophone L1 or non-school purposes. These results suggest that none of our classrooms correspond to a prototypic Type A program (i.e., one in which only analytic activities are used), or to a prototypic Type E program (one in which only experiential activities are used). The classrooms in our sample fall somewhere in between the two extremes.

Colt Part II

Part II of the COLT observation scheme analyses the pedagogic features of verbal interaction during classroom activities. Each subcategory in Part II was calculated twice, once as a percentage of total observed time, and once as a proportion of the superordinate category. Taking teacher's use of French as an example, we note that Type E teachers used the target language for 49.81% of total observed time (Table 8A, p. 106), but they used L2 for 96% of the time they were actually speaking, i.e., when time was calculated as a proportion of the superordinate category (Table 9A, p. 110). The total time coded for Part II activities, the total number of speech turns which occurred during the one-minute coding periods, and the proportion of teacher to student turns for each school, are presented in Table 10 (page 114).

The data for teacher and student interaction, with subcategories calculated as a percentage of total observed time, are presented in Tables 8A-L (pp. 106 - 109). When Part II categories were calculated as a percentage, significant differences between Type A and Type E classrooms were found in terms of student sustained speech, student reaction to message/code, and student topic incorporation. As indicated in the tables, students in Type E classrooms spent a greater amount of time producing sustained speech than students in Type A classrooms (18.14% and 3.87% respectively). This difference was significant ($p < .05$). Type E students spent significantly more time reacting to the message (6.92% compared with 1.88% for Type A students, $p < .01$). Also, topic expansion by students occurred significantly more often in Type E than in Type A classrooms (2.41% and 0.82% respectively, $p < .01$).

In Type E classrooms a significantly greater percentage of teacher and student talk was coded as unintelligible. In Type E classrooms 3.79% of teacher talk was unintelligible compared with 0.80% in Type A classrooms ($p < .05$), and 11.99% of student talk was unintelligible compared with 6.97% in Type A classrooms ($p < .05$).

No statistically significant differences were found in use of target language, information gap, discourse initiation, or form restriction when Part II categories were calculated as a percentage of observed time.

The data for teacher and student interaction, with subcategories calculated as a proportion of the superordinate category, are presented in Tables 9A-L (pp. 110 -113). When Part II categories were calculated as a proportion, significant differences were found between Type A and Type E classrooms in terms of student sustained speech, student form restriction, and teacher and student reaction to message/code.

As indicated in the tables, the proportion of ultraminimal turns was similar in the two types of classroom, but Type A students made significantly more minimal turns (51% compared to 32%, $p < .05$), and Type E students made significantly more sustained turns (26% compared to 8%, $p < .01$). In Type E classrooms the choice of linguistic item used by students was less likely to be restricted (47% of student utterances compared with 81% in Type A classrooms, $p < .01$). In Type A classrooms there was a greater likelihood that students would react to the code rather than the message (34% compared with 13% in Type E classrooms, $p < .05$). Reactions to the message in Type E teachers' speech were significantly more frequent than in Type A teachers' speech (76% and 35% respectively, $p < .01$).

No statistically significant differences were found in use of target language, information gap, teacher sustained speech, teacher topic incorporation, or student discourse initiation when Part II categories were calculated as a proportion of the superordinate category. As indicated in the tables, teacher and student verbal interaction were almost always in the target language in both types of classroom. On the other hand, the feature Discourse initiation occurred extremely rarely in student speech in all the classrooms observed. Students in Type E classrooms had a greater tendency to give unpredictable information, and to make genuine requests, than students in Type A classrooms, although the differences were not significant.

To summarize, an analysis of the observational data revealed significant differences between Type A and Type E classrooms with regard to: topic control by teacher and student, student extended and minimal written text, teacher and student reaction to message/code, student sustained and minimal speech, form restriction, source/purpose of materials, and student expansion. Type A classrooms made significantly more use than Type E classrooms of the following features (S = students):

topic control by teacher
minimal written text (S)
minimal utterance in spoken interaction (S)
reaction to code rather than message (S)
restricted choice of linguistic item (S)

Type E classrooms made significantly more use than Type A classrooms of the following features (T = teacher, S = students):

topic control by student
extended written text (S)
sustained speech in spoken interaction (S)
reaction to message rather than code (T, S)
topic expansion (S)
use of student-made materials

Moreover, in Type E classrooms, a significantly greater percentage of student and teacher speech was coded as unintelligible.

These results show that our first two assumptions (p. 56) were correct: student and teacher behaviour did vary from class to class within the sample, and the differences were characterisable in terms of at least some of the COLT categories. We found significant differences both in the types of activities and in the amount of time devoted to each type, and these differences enabled us to rank classes along an experiential-analytic scale. Furthermore, the COLT analysis suggests that the sample as a whole is intermediate along an absolute scale ranging from prototypically high-experiential classes at one extreme, and prototypically high-analytic classes at the other end of the scale.

5:3 Teacher Questionnaire

As indicated in section 3:4, a teacher questionnaire was used to obtain information about teaching/learning activities throughout the year, including those occasions when an observer was not present. The concept of experiential versus analytic activities incorporated in the questionnaire derived partly from a review of the communicative language teaching literature, and partly from classroom observations carried out during the development and validation studies. Teachers were presented with a list of activities and asked to indicate whether each activity was performed 'never', 'rarely', 'sometimes', 'quite often' or 'very often' during the year. Table 11 (see p. 115) contains a list of experiential activities which were performed 'quite often' or 'very often' during the year in Type A and Type E classrooms. Table 12 (p. 116) contains a similar list of analytic activities.

A comparison of the tables shows that there was considerable overlap of activities in the two types of classroom. At the same time, however, there was a tendency for experiential activities to occur more frequently in Type E classrooms, and for analytic activities to occur more frequently in Type A classrooms. For example, teachers indicated that the following experiential activities occurred more often in Type E than in Type A classrooms during the year: students practise conversational skills by talking in pairs or groups; students listen to spoken French materials that are not specifically produced for FSL learners; students are taught aspects of paragraph and text structure; the teacher focuses on stylistic appropriateness and/or logical organization of text in correcting written work; creative writing tasks are assigned for homework.

Teachers also indicated that the following analytic activities occurred more often in Type A than in Type E classrooms during the year: students do single-sentence, fill-in-the blank exercises; students do guided writing tasks based on pictures, diagrams, etc.; students practise by repeating words and/or sentences after hearing teacher/tape, or by doing oral substitution or transformation exercises; students listen to extended spoken texts, specially recorded for FSL learners and delivered at a reduced speed, carefully articulated, etc. However, it is interesting to note that the following analytic activities were noted as having occurred 'quite often' or 'very often' during the year by all the teachers in the sample: students do grammar and/or vocabulary exercises either orally or in written form; the teacher focuses on spelling, grammar and/or use of vocabulary in correcting written work (i.e., as distinct from focusing on stylistic appropriateness or logical organization of text); the teacher focuses on pronunciation, grammar, and/or use of vocabulary in correcting oral work (i.e., rather than focusing on the message being conveyed); the teacher assigns grammar and vocabulary exercises in a single-sentence format for homework.

All the teachers in the sample, with one exception, used a basic, structurally-graded textbook (e.g., *Vive le Français* or *Passeport Français*). In addition, the teachers used a variety of supplementary material, including videotapes, records, newspaper and magazine articles, songs, stories and poems. One Type E teacher listed "guest speakers developing relevant topics and discussion with students" and "restaurant and field trips during class or after class time" in the list of supplementary activities. None of the Type A teachers mentioned these types of activity. The questionnaire also showed that all the teachers in the sample made some effort to use French with the students before and after the actual French period, and encouraged students to use French during class for administrative and classroom management purposes.

These results confirm the findings of the COLT analysis, namely, that both Type A and Type E classrooms had analytic characteristics (e.g., focus on form, use of teacher-centred activities, a general lack of student-initiated discourse). At the same time, there were numerous differences in pedagogic orientation between the two groups, as confirmed by the results of the teacher questionnaire.

5:4 Classroom Profiles

In addition to the COLT data and the teacher questionnaires, a descriptive profile was compiled for each of the classrooms on the basis of field notes made during the observation periods. Although the profiles could not be quantified, they enabled us to record details which would otherwise be difficult to capture, and to obtain a broad picture of individual teaching styles. For example, the following Type A profile is reminiscent of highly controlled, teacher-dominated classrooms in which students spend most of their time focusing on form with little opportunity for spontaneous interaction:

Class 8 profile (Type A)

Throughout the four observations the teacher followed the same pattern: first, a review of vocabulary items learned in the previous class, then a review of the grammatical rule from the previous class. The students write their sentences on the blackboard and the teacher corrects. After this the teacher corrects homework; usually fill-in-the-blank exercises requiring use of the correct verb form, etc. Again the students write their answers on the blackboard and the teacher corrects. Up to this point there is relatively little opportunity for students to use whatever forms they have learned in a communicative context. The last third of the class is devoted to reading comprehension. The teacher gives the students a handout consisting of a short reading passage with content questions. The text is never really discussed until after the students have answered the content questions. These discussions are controlled by the teacher. This is a very structured class, basically always following the same steps. The teacher corrects gender, spelling and pronunciation, and chooses all the topics for discussion.

The following Type E profile describes a more experiential classroom environment in which students prepare their own materials, discuss topics of genuine interest, and regularly engage in extended spoken and written discourse:

Class 2 profile (Type E)

This teacher has no formal text. At the beginning of the year he gives each student a copy of a guide focusing on improving oral communication, and a copy of a guide for improving writing skills, both of which he has prepared himself. The students are asked to write two major essays during the year, on topics related to Le Petit Prince, a novel the teacher has selected. Over a period of two months the book is read and discussed on many levels — literary, religious, philosophical, etc. The students are also asked to write one composition a week. Grammatical errors arising out of the compositions are discussed, and the teacher prepares handouts to practise points where the students are having difficulty. There is a class magazine in which the best compositions and poems are published. This journal appears twice a year and the students are very enthusiastic about it. The students have ample opportunity to talk about anything they choose. The teacher corrects pronunciation, verb tenses, gender, etc., but he usually waits until the student has finished speaking before making the correction.

As a result of examining the observational data, the teacher questionnaires and the classroom profiles, we concluded that there were major differences between the two types of classroom which could have an effect on learning outcome. Although Type A classes included a wide range of activities, work in these classrooms tended to be dominated by the requirements of a predesigned, linguistically organized syllabus. Type E classes, on the other hand, were somewhat more flexibly organized, so that selection of content, choice of speech act, distribution of roles, etc., had a greater chance to arise spontaneously out of on-going classroom work or in response to the students' immediate on-the-spot needs and desires.

As previously discussed, neither class in the Type E group corresponds to a prototypically high-experiential class (i.e., one where all and only experiential activities are used), and none of the classes in the Type A group corresponds to a prototypically high-analytic class (one where all and only analytic activities are used). Nevertheless, the significant differences between the two groups should permit us to examine the extent to which they contribute to differences in student knowledge and performance.

6. PROFICIENCY PREDICTIONS

In this section, we will define how the tests operationalize the theoretical constructs at issue and how classroom activities were expected to relate to language proficiency.

We had assumed initially that students whose classes were defined as analytic by the COLT (Type A classes) would in fact get more explicit instruction about grammatical form. We had also assumed that analytic classes would get more explicit correction of errors of form, since accurate production would be deemed to be a priority.

We had assumed that these same students would get little or no instruction on the organization of discourse, sociolinguistic information or strategies for speaking. In contrast, we had expected that experiential (Type E) classes would spend time on such activities and would, in addition, get practice in using the language for communicating meaning. This, in itself, would entail more exposure to extended spoken and written language.

Given the additional, and for our purposes necessary, assumption that there is a direct connection between what core French students do in class and the nature and extent of their L2 proficiency, it seemed likely that students from analytic and experiential classes would know different things about French and would be differentially skilled in using their knowledge.

We predicted on the basis of these assumptions:

- **Hypothesis 1A** Students from the analytic classes would score significantly higher on the multiple-choice grammar test than students from the experiential classes.

- **Hypothesis 1B** Students from the analytic classes would score significantly higher on the oral grammatical measures.

These two predictions follow naturally from the assumption that students in analytic classes know more about grammar and form because they spend more time in classroom activities focused on them. Nevertheless, the predictions are by no means trivial. Krashen (1982), for example, disputes the claim that explicit teaching about grammar affects the learner's tacit system of grammatical knowledge or production skills. Consequently, if Krashen's distinctions between learning and acquisition are correct, one might expect, contrary to our predictions, that the students in experiential classes would do as well as the students in analytic classes on oral grammar measures. They would not necessarily do as well on the written test of discrete points of grammar.

- **Hypothesis 2A** Students from the experiential classes would score significantly higher on the discourse measures of the note and the letter.

- **Hypothesis 2B** Students from the experiential classes would score significantly higher on the oral discourse measures.

Once again, these predictions follow naturally from the assumption that students whose classes are classified as experiential by the COLT would actually spend more time in activities whose focus is on the organization of discourse, on producing coherent text, on defining and manipulating cohesive features of French, etc. Furthermore, we had assumed initially that students whose classes were defined as experiential would actually produce extended discourse significantly more often than the analytic classes.

- **Hypothesis 3A** Differences between the sociolinguistic scores on the formal letter and the informal note would be significantly greater in the case of the experiential classes than they would be in the case of the analytic classes.

- **Hypothesis 3B** Differences between the sociolinguistic scores on the oral interviews would be significantly greater in the case of the experiential classes.

These hypotheses depend upon the assumption that students in classrooms defined as experiential would actually spend more time on activities differentiating formal and informal varieties of French, and on activities which identify characteristics of speech styles, that they would get more practice at producing speech acts in extended discourse, etc.

- **Hypothesis 4A** Students from the experiential classes would score significantly higher on the listening comprehension test.

- **Hypothesis 4B** Students from the experiential classes would show greater comprehension of questions asked on the oral interview as revealed by strategic proficiency measures encoding noncomprehension of the questions posed.

These two hypotheses make sense if we assume that classes identified as experiential actually spend more time listening to naturalistic speech, manifesting all of the specific phonological features of discourse (reduced vowels in unstressed syllables, varying rhythms, contentful intonation patterns, focus stress, etc.), as well as other features of discourse (repetition of words, use of morphologically related words, use of hyponyms, synonyms, etc.). These properties can only be manifested over extended text, in sequences of turns in dialogue, etc. Furthermore, it is plausible to assume that students who spend time listening to naturalistic speech develop strategies for deriving meaning globally from the discourse. The extent to which students listen to naturalistic discourse cannot be determined directly from the COLT but would be inferable from those categories related to the use of extended text and the nature of student/teacher or student/student interactions. We therefore made the assumption that if experiential classes spend significantly more time listening to extended text then they also spend significantly more time listening to naturalistic speech.[12]

- **Hypothesis 5** Students from the experiential classes would show higher accuracy rates on the oral sentence repetition task.

The sentence repetition task requires various kinds of knowledge and skills. In order to repeat the specific sentences of the text, the subject must be able to parse the sentence (and not just listen for gist). Parsing requires grammatical information since the student must be able to match the incoming signal against representations of words previously stored in long term memory, identify the sounds as belonging to specific units, group the units and identify a meaning. Nonetheless, knowledge of grammatical structure or generalizations is not sufficient. Students must be able to recognize words even when they are subject to discourse phonology, and they are not familiar with the accent of the speaker, etc. Consequently, we felt that students who are used to listening to discourse and who have developed strategies for decoding the sound input would be favoured.

7. PROFICIENCY TESTS

Tests were designed to provide measures of grammatical, discourse and sociolinguistic competence, and listening skills in French.[13] As procedures varied with each instrument, details will be discussed in relation to each test.

7:1 The Multiple Choice Grammar Test

The multiple choice grammar test consisted of 38 items, all of which had three possible options as a response.[14] Items were scored by computer as right or wrong. The proportion of correct responses out of 38 provided the total score, which was adopted as the measure of grammatical competence for each class as a whole.[15]

Among the questions on the multiple choice grammar test were several which bear on the subcategorization and selectional properties of verbs. For example, question 12 requires that the student decide what lexical item must follow the verb s'attendre:

12. Je m'attends ____________________ ce qu'elle arrive bientôt.
    a) à
    b) pour
    c) que

Learning the word attendre entails learning various grammatical properties particular to this word (that it is transitive, that it belongs to the class of 'reflexive' verbs, and that it can be followed by a sentential complement). Among these properties is the fact that the verb must be followed by the preposition à. Thus, the verb is subcategorized for this preposition. Questions 1, 12-13, 18, 20, 28-29, 35, 37-38 were all structured to elicit this kind of information.

Questions 2, 23, 25, 30 were designed to elicit information about gender. As is well known, French nouns are all associated with either masculine or feminine gender. This association is arbitrary insofar as it does not encode any semantic or phonological information, and must be learned for a particular noun or derivational suffix. An example is given below:

29. Les États-Unis se trouvent entre le Canada et ________ Mexique.

a) la
b) le
c) --

While in principle a distinction must be drawn between the assignment of gender to derivationally complex words (those bearing derivational suffixes such as imitation, atterrisage, etc.) and words that are morphologically unanalysable (e.g. table, oiseau, mal), our questions made no attempt to differentiate the two types of words.[16]

Questions 7, 9, 14, 17, 19 and 21 were designed to elicit information about agreement processes such as agreement of the past participle with the subject, agreement of the adjective with a noun in a noun phrase or in a predicative clause involving être (of the type NP être NP). For example:

7. Marie était bien ____________________ aujourd'hui.
a) habiller
b) habillée
c) habillé

Question 7 illustrates a context where the predicative adjective must agree with the subject. In these questions, the student is not required to distinguish between different phonological forms of the word but rather to pick out that orthographic form which encodes the agreement process.

Subtests were done on the subcategorization, gender and agreement questions to permit a comparison with the grammatical measures of the oral interview (which will be discussed below). The proportion of correct responses out of the total number of responses provided the measure of proficiency for each subtest.

Reliability coefficients for the pre-test and the post-test were computed separately. The Guttman split-half coefficient was .73 on the pre-test and .69 on the post-test, indicating that the grammar test was reasonably reliable.

7:2 The Letter and Note

The writing tasks were designed to elicit information about sociolinguistic competence, that is to say, they were designed to provide a task where students would be required to formulate a given speech act (a request, a command or a threat), varying in the use of politeness markers according to the nature of the relationship between the student (as writer) and the addressee. In the case of the letter, the student had to make a written request, and had to understand that the situation (discourse with a higher status, unknown adult) required the use of a formal and polite register. Such a register would be realized by the use of various politeness expressions and formal attenuators. In the case of the note, the student had to adopt the role of his/her own parent, taking on a position of authority vis-à-vis the student and requesting, commanding or threatening him or her to tidy the house. In this task, the writing would be directed to a lower status, younger intimate.

Sociolinguistic scoring

The letter and note were scored for the presence or absence of seven attenuators. If the forms were present, a score of 2 was assigned, otherwise a score of 1 was assigned. The particular forms scored for were (a) the conditional form of the verb, (b) modal verbs, (c) interrogative structures, (d) politeness expressions such as s'il vous plaît, (e) vocabulary appropriate to formal registers. We also scored for the presence of polite second person forms of the pronoun vous. Here a three-point scale was used: a score of 3 was assigned if most or all of the pronouns were vous; a score of 2 was assigned if there was a mixture of vous and tu; a score of 1 was assigned if there was no use of vous. The formal closing was also rated on a three-point scale: 3 was given for a very formal closing, 2 for a moderately formal closing and 1 for an informal closing. For further discussion of these features, see the Year 2 Report, pp. 29-31.

As in previous studies, a difference score was calculated on the assumption that formal text would be characterized by the presence of some or all of the seven features. An ideal score for the written letter would thus be a high score (a maximum of 16). An informal text, it was hypothesized, would have few or none of the attenuators. Thus, an ideal score for the note would be low (a minimum of 7). The score for sociolinguistic competence involved subtracting the score for the note from the result for the letter. The ideal score was 9 (16 minus 7) and meant that the student had maximized the number of attenuators on the letter and minimized their use on the note.

Scoring of the letter was done by scorer 1 (a native speaker of French). A random sample of approximately 10% of the letters was scored again by scorer 3 (a fluently bilingual anglophone). The two scorers agreed more than 90% of the time on all scores. Mean differences on scoring ranged from .01 to .06. The same procedures were followed for the note with similar results. Mean differences on scores ranged from .01 to .05.

Discourse scoring

The letter and note were rescored for aspects of discourse. We were interested in establishing the degree to which the students could produce both coherent and cohesive text (Halliday and Hasan 1976, Carrell 1982). Tests were scored for (a) the presence of sufficient background information, (b) task fulfillment, (c) consistent use of text features.

Providing sufficient information included the appropriate identification of referents to provide sufficient context for the successful interpretation of the speech act. In the case of the letter, where the student did not know the addressee, the situation demanded that the writer presuppose a minimal amount of shared information with the addressee. Consequently, the student had to identify and locate all participants and relevant objects in making the request. In the case of the note, the student might presuppose a great deal of shared information and so the note could be considerably more terse. On each test, there were five pieces of information to be identified.

Task fulfillment included the requirements that the request, command or threat be successfully made, that the text follow some logical ordering, and that a text organization appropriate to the discourse genre (e.g., openings and closings for letters and notes) be used. Logical ordering of the text was only scored when all of the relevant features were present, namely opening, request, rationale and closing.[17]

In scoring for consistency of text features, we examined the appropriate use of anaphora across sentences, consistent use of French, and consistent use of either _tu_ or _vous_ through the text, i.e., three features in all. In the sociolinguistic scoring, we evaluated the situation-appropriateness of _tu_ and _vous_. However, in the discourse scoring we took a somewhat different approach. Here we ignored whether or not the particular choice was suitable and focused on whether the student made a choice and stuck to it. The assumption was that consistent use of a single form of the pronoun would produce more cohesive (and presumably more coherent) text than random use of _tu_ and _vous_. Similar remarks can be made about the use of pronouns and noun phrases. A comparison of the following examples will make this point clear:

(a) J'ai vu la bicyclette dans son garage. La bicyclette est belle. Je veux utiliser la bicyclette.

(b) J'ai vu la bicyclette dans ton garage. Elle est belle. Je veux l'utiliser.

In (a) the writer repeats the full noun phrase _la bicyclette_ instead of using a pronoun. The first example also contains the erroneous use of _son_ instead of _ton_. Neither of these errors occurs in (b). Therefore, (a) is more cohesive and coherent than (b), and would receive a higher score.

We assumed that a text which was entirely written in French, with no codeswitching and no borrowing, would be more coherent and cohesive. Thus we scored for the presence or absence of these features.

In all cases, scorers looked simply for the presence or absence of the feature. No attempt was made to differentiate degrees of correct use. Thus, we did not distinguish on this test between students who used the pronoun _tu_ in three-quarters of the possible contexts and those who used it in one-quarter of the contexts. Both groups of students were scored as using the features inconsistently.

Scoring was done by scorers 2 and 3 (both bilingual anglophones) who consulted frequently. Each person scored about half of the letters. A random selection of twenty letters was then given to the other scorer to be analysed. Scorers agreed better than 85% of the time on their ratings. Mean differences ranged from .01 to .07. Using the same procedures on the notes, scorers agreed more than 85% of the time. Mean differences in results ranged from .01 to .06.

7:3 Listening Comprehension

The listening comprehension test (LCT) consisted of a series of recorded texts involving different accents and different types of language. The texts were drawn from two different sources, namely the _Test de compréhension auditive_ from the Bilingual Education Project, and the _Test of Listening Comprehension_ from the _IEA: Population 4_ (see Carroll 1975). The BEP test was originally developed for immersion classes. The questions were altered somewhat so that the students would not be looking for discrete points of information but rather would be able to respond correctly on the basis of a global comprehension of the text. The recorded texts were checked for authenticity by a native speaker of French. The LCT was piloted with a group of sixteen-year-old core French students in the early fall prior to administration of the pre-test. The test consisted of 19 multiple choice items, each with four possible answers for the response.

Items were scored as either right or wrong and the proportion correct provided the total score for listening comprehension.

As with the grammar test, pre- and post-test results were subject to a reliability test. The Guttman split-half coefficient was .52 for the pre-test and .67 for the post-test, indicating that the test was not highly reliable. Two items in particular correlated poorly, namely items 6 and 7. Scores were recalculated deleting these questions. The results were not much different; a coefficient of .54 was obtained for the pre-test and .63 for the post-test. Since the reliability was not significantly improved by deleting items 6 and 7, all items were used in computing the LCT scores.

7:4 Oral Interview

The interview was semi-guided. As with the written work, the interview was designed to elicit information about the students' grammatical, sociolinguistic and discourse competence. Different questions focused on specific aspects of competence, as indicated below.

Grammatical scoring

Some questions on the oral interview were designed to elicit verb tense usage, in particular the present tense, the periphrastic future, the conditional, the imperfect and past perfect tenses (questions 2-5, 9, 13 and 15). Other questions were designed to elicit the use of prepositions and stative verbs (question 12). A separate question (number 16) focused on verbs of direction and locomotion.[18] These and other questions would also provide material for further grammatical analysis, namely subcategorisation and selection, agreement, auxiliary selection, and gender. Scoring was done in all cases by counting the total number of correct and incorrect uses of a form out of the total number of obligatory contexts.

Scoring of the oral grammatical measures was done by one individual (scorer 2) on all measures except the subject-verb agreement measures and the verb tense measures. Those questions were scored by a fluent non-native speaker of French (scorer 4). In order to check the reliability of the scoring, a sample of four students was selected for rescoring by scorer 2. The two scorers were in agreement some 95% of the time on the 240 data points that were checked, indicating that the scoring was quite reliable.

Sociolinguistic scoring

A measure of sociolinguistic proficiency was based on responses to questions 17 and 18 of the oral interview. These questions were designed to elicit requests, commands or warnings to same status intimates and differing status non-intimates. So, for example, one question involved making a request to a fellow student and then making the same request to the principal of the school. The other question involved giving a warning. In formulating the warning, students had to address themselves to a friend of the same age, and then to an elderly woman.

We anticipated eliciting basically the same kinds of attenuators that we scored for on the letter and note. Thus we scored for the use of address terms, polite singular referent vous, politeness expressions, modal verbs, question forms, conditional tense, and use of an explanation to justify making the request or warning, i.e., eight features in all.

A point was assigned for each of the features. Optimal scores were high in the formal context (a maximum of 8) and low in the informal context (a minimum of zero). A difference score was arrived at by subtracting the low score from the high score as the measure of sociolinguistic proficiency.

Scoring of the oral sociolinguistic measures was done by scorer 3. A sample of eight students was rescored by scorer 4 in order to check reliability. The scoring of questions 17B pre-test and 18A post-test proved unreliable, with alphas of less than .47. The other scorings were very reliable with alphas ranging from .89 to .97. The problem with the two unreliable scorings was traced to counts of the presence or absence of vous. The two scorers consulted on the scoring criteria for this feature and rescored. The second scoring proved reliable.

Discourse scoring

It was expected that some questions on the oral interview, i.e. those eliciting verb sequences, would also provide material for an assessment of discourse competence. We selected consistency of verb tense usage (i.e., present, past, periphrastic future and imperfect) as one feature of discourse proficiency. We also examined the consistency of pronoun (tu/vous) usage on those questions (namely 17 and 18) which were used to obtain information about sociolinguistic proficiency. On both of these counts, a consistency score was assigned only if there were two or more examples of the relevant form. Thus, if a student used the present tense only once in responding to a question, then no consistency score was assigned. Similarly, if the student used a pronoun only once in expressing a given speech act, then no consistency score was assigned. In addition, a task fulfillment score was obtained for each speech act on questions 17 and 18. The total discourse score was the sum of these features.

Scoring of the oral discourse measures was done by scorer 3, and a sample of eight students was rescored by scorer 4. The scorers were in agreement on task fulfillment with alphas ranging from .73 to 1.0. The scorers agreed well on the consistent use feature with an alpha of .94 for all verb tenses and pronouns.

Indices of non-comprehension

To capture the extent to which students understood questions asked during the oral interview, we analysed the entire interview for indices of noncomprehension. Various categories were established: question misinterpretation (the student responded automatically but not to the question asked); unsolicited question repetition (the question was repeated because the interviewer felt, because of a hesitation in responding or because of kinetic clues, that the student had not understood);[19] direct requests from the student for a repetition of the question; direct requests for an explanation, a translation of the question or some part of it; indirect requests for repetition or an explanation made by asserting noncomprehension (e.g. I don't understand) or an inability to do the task (e.g. I don't know); an English response to the question; requests for confirmation (indicated by the use of rising intonation in utterances which were not questions). The total non-comprehension score was a sum of the presence of the above features.

Scoring of the indices of noncomprehension was done by scorer 3, and a sample of eight students was rescored by scorer 4. The scorers agreed well on the various features, with alphas in the range .78 to .97.

Sentence repetition task

To obtain a measure of the students' ability to parse French text (and to contrast with the comprehension measures of the LCT) a sentence repetition task was devised and administered at the end of the oral interview to one randomly selected student from each class. It consisted of the following text recorded by a native speaker of French:[20]

> Six‿étudiants sont‿en train de manger dans‿un restaurant. Ils‿ont très‿envie d'avoir des desserts et la nourriture disparaît vite. Le serveur vient‿à la table et il leur annonce: "Je regrette mais on‿a plus de dessert. Je n'peux pas vous‿ apporter ce que vous‿avez commandé."

The text was preceded by instructions and by two examples illustrating a stimulus sentence and then its repetition. After hearing the whole text, students heard each sentence individually. After each sentence, they were required to repeat what they could.

As can be seen, the sentences formed a text. It was felt that this would produce more natural language. Each sentence was long enough (approximately 15 syllables) to require processing for meaning. In other words, students would not be able to retain the stimulus in short-term memory and then repeat without understanding the message. The form of the stimulus was designed to elicit liaison, pronunciation of rounded vowels, and nasal vowels. We had hoped to study these features of pronunciation. This type of analysis, however, proved to be impossible since the repetition task was, by and large, too difficult for the students. They simply could not parse the sentences.

Two different scores were calculated for each sentence. The first calculation was scalar. A score of 1 was assigned if no repetition was attempted; 2 was assigned if there was a partial but inaccurate repetition of any part of the sentence; 3 if there was a partial but accurate repetition of at least a syntactic phrase in the sentence (usually the last one of the stimulus); 4 if there was either an incomplete but basically accurate reproduction of the sentence as a whole or a full but slightly inaccurate repetition of the sentence; 5 if there was a full and accurate repetition of the stimulus. The second score consisted of a sum of the number of syllables accurately reproduced for each sentence. There was no requirement that the syllables should compose a syntactic constituent (i.e., that they should form a word, phrase etc.).

Scoring was done by scorers 3 and 4. As indicated above, the scorers rated the students' repetition of each of four sentences on a scale of 1 to 5. High inter-rater reliability was obtained (alpha .97 on both pre-test and post-test).

8. PROCESS/PRODUCT FINDINGS

8:1 Overview

The analysis of the effects of classroom process on second language proficiency proceeded in three stages. First, the eight classrooms were divided into two groups according to the overall COLT score for experiential vs. analytic orientation, as discussed in section 5:1. Those students who participated in relatively experiential programs (classrooms 2 and 5) were pooled together to make up Group E, while those who participated in relatively analytic programs (classrooms 1, 3, 4, 6, 7 and 8) constituted

Group A. The relative improvement of the two groups on the various proficiency measures was examined by analysis of covariance, using pre-test scores as the covariate, and post-test scores as the dependent variable.[21] Analyses were carried out using the MANOVA procedure in the Statistical Package for the Social Sciences, version X (SPSS-X).

As discussed below, the results of the first stage showed few significant differences between experiential and analytic groups. At the second stage, the same analysis was repeated using only the two most extreme classrooms from each end of the continuum, in order to maximize the chances for differences to emerge. The second analysis compared students from the two most experiential classrooms (Group E - classrooms 2 and 5) to those from the two most analytic classrooms (Group A* - classrooms 4 and 8), again using analysis of covariance.

At the third stage, the correlations of each COLT observation category with adjusted class means on each of the post-test measures were examined, using the classroom as the unit of observation. These correlations were interpreted to indicate which particular features of the classroom process, as identified by the COLT, were beneficial or non-beneficial to learning.

8:2 Comparison of Experiential and Analytic Groups (Groups E and A)

Pre-test and post-test performance levels and adjusted post-test means for Group E and Group A on each of the proficiency measures are presented in Table 13 (p. 117). These means are expressed as percentages of the total score possible. The adjusted post-test means are statistically adjusted for pre-test differences, and give an indication of the post-test score that each group would have obtained had they started out equal. For the written measures, means are shown both for the total sample in each group and for the subsample of each group that were interviewed. For the oral measures, of course, means are available only for the interviewed subsample.

The most striking thing about these data is the lack of difference between Groups E and A. None of the differences between groups on adjusted post-test scores is significant, although the difference in favour of Group A on the grammatical multiple choice written test is nearly significant ($p < .06$) for the whole group. For the interviewed subsample the difference also favours Group A, although it is not significant. For the other written tests, the differences between groups are not significant, and in two of the three cases, in different directions for the whole sample and the interviewed subsample.

For the oral tests, Group E does marginally better in grammar — in contrast to the written grammar test where Group A does better. However, none of the differences between oral tests is significant, and therefore no conclusions can be drawn from these results. Group E gets somewhat better results on the repetition task, but again, this difference is not significant.

Three sub-scales were defined on the grammatical multiple choice written test by selecting items that dealt with subcategorization, gender and agreement respectively (see section 7:1). Similar sub-scales were constructed for the grammatical oral test. Pre-test and post-test performance levels and adjusted post-test means for these sub-tests are shown in Table 14 (p. 118). The advantage of Group A on the written grammar test is seen to be largely due to the agreement sub-scale. For the oral test, Group E showed more improvement on the gender sub-scale but less improvement on the other two sub-scales. However, again, none of these differences is significant, and no conclusions can be drawn.

The sociolinguistic written scores presented in Table 13 were constructed as difference scores between performance on a formal letter and on an informal note (see section 7:2). The individual scores on letter and note are presented in Table 14 (p. 118). The difference between Group E and Group A is significant in the whole sample for both letter and note. In both cases Group A has the higher score, indicating that students in this group used a greater number of formal markers. However, neither group is better than the other at signalling the distinction between formal and informal, as shown by the non-significant difference between groups on written sociolinguistic difference scores in Table 13. These results suggest that Group A has more knowledge of the formal markers that are necessary to achieve an appropriate formal tone in the letter, while Group E may have more control over the informal style which is appropriate for a note. Neither group appears to master the distinctions which the forms are intended to convey. Differences in the interviewed subsample are in the same direction, but are not significant.

Scores for discourse on letter and note (Table 14) are consistent in showing that Group A has improved more than Group E, but the differences are not significant.

Perhaps the most interesting thing about the results for Groups E and A in Tables 13 and 14 is not the differences that emerge -- possibly in the grammatical written test, and in the letter and note, -- but the lack of differences between the groups in most aspects of language acquisition that were tested. It should be stressed that this lack of difference exists despite clear differences between groups on classroom process variables that distinguish more experiential from more analytic teaching methods, as discussed in section 5:2. Also, the lack of difference exists despite clear and consistent differences between individual classrooms in improvement on the proficiency measures. These classroom differences in French proficiency will be discussed in section 8:4 below. The lack of significant differences between Group E and Group A shows that real differences in overall classroom environment do not necessarily translate into significant differences in proficiency.

At the same time, it must be remembered that the range between experiential and analytic in the present study is restricted by the fact that all our classrooms were fairly analytic relative to programs such as French immersion (Fröhlich et al. 1985). If similar studies were conducted with other classes within core French or with other types of program, it is possible that the range between the most experiential and the most analytic classes would be greater, and that there would be more differences in outcome between the two orientations.

We might interpret this failure of correspondence in several ways. The most extreme interpretation would be to conclude that the environment has no effect on the learner's knowledge and proficiency. Since there are strong reasons to suppose that this conclusion is false, we set it aside. A second possibility is to suppose that the observed similarities among the classrooms had more impact on the proficiency of students than the observed differences. It will be recalled that with regard to specific features both groups were fairly analytic. It may be the case, then, that an explicit focus on form and the teacher-centred activities of both types of classroom in themselves explain the lack of difference between Groups A and E on our proficiency measures. Only further research will clarify this issue.

Our present interest, however, is the range of difference that does exist within the core French program. In order to give the best chance within core French for differences between experiential and analytic orientations to emerge, we considered a new grouping of A-type students, Group A*, which included only those students in the two most analytic classrooms, 4 and 8. This allowed us to carry out analyses similar to those just presented, but between the two most extreme classrooms at each end of the continuum, i.e., between Group E (defined as before) and Group A*.

8:3 Comparison of Experiential and Analytic Groups (Groups E and A*).

Pre-test, post-test, and adjusted post-test mean performance levels for Groups E and A* comparable to those shown in Table 13 were examined.[22] On the grammatical multiple choice written test, Groups E and A* were now significantly different ($p < .05$), with Group A* showing more improvement than Group E. The other differences were not significant. That is to say, the differences on the LCT, the sociolinguistic and discourse scores for note and letter, and the oral measures did not distinguish the two groups. In sum, the results for Groups E and A* were therefore the same as the results for Groups E and A, except that the difference on the written multiple choice grammar test was now significant.

Detailed scores comparable to those shown in Table 14 were also examined. Again the difference between groups on the sociolinguistic scores for letter and note were significant, with Group A* showing more use of formal markers in both cases. In addition, the difference in favour of the analytic group on the written grammar subtest for agreement was now significant ($p < .05$), although not for the interviewed sub-sample. In general, then, there was little difference between the results for Groups A and A*, although the use of only the extreme groups in A* did increase the significance in the written grammar test, the use of formal markers on the note and letter, and use of agreement.

We also examined, in the context of Groups E and A*, the sub-scores that make up the sociolinguistic and discourse tests. First, with regard to the sociolinguistic sub-scores, Group A* scored significantly higher than Group E on the use of the conditional verb tense in the letter ($p < .001$). The mean scores show that this is not because Group A* used conditionals extensively, but because Group E hardly used conditionals at all, a fact that is in agreement with the results on the grammatical tests, where neither group showed mastery of conditionals. There were no significant differences between groups on the other sub-scores of the sociolinguistic written test, either in the letter or the note, including use of formal vocabulary, modals, politeness expressions, formal second person pronouns, and indirect questions. Mean scores show that students made reasonable use of indirect questions and politeness expressions in the letter, but that there was a complete absence of formal vocabulary.

The sociolinguistic oral test showed no significant differences between groups on any of the sub-scores, and showed very limited use of conditionals, modals, direct and indirect questions, and formal vocabulary. Explanation was the attenuator most frequently used, followed by politeness expressions. Students tended to use vous rather than tu on all the sociolinguistic questions.

With regard to the written discourse sub-scores, there were no significant differences in making the request in the letter, nor in making the request/command/threat in the note. Group A* scored significantly better (p .01) in providing a rationale for the letter, but not for the note. This suggests that students in the more analytic classrooms were better able to perceive the need for a reason to make requests of an unknown addressee. It may be that students in general feel that parents are not required to explain commands concerning duties about the house. There were no significant differences in the logical ordering of the information on either the letter or note, which was adequate among the small number of students who selected all of the relevant features and could therefore be scored. Group A* were also better at providing a closing in the letter (indicating knowledge of the structure of genre) ($p < .01$); there were no significant differences in providing an opening or closing in the note. There were no significant differences on the cohesion scores (consistent use of second person pronoun; appropriate use of anaphora; and consistent use of L2), nor on task fulfillment, on either the letter or note. The mean scores indicate very little use of anaphora. Also, there were no significant differences between groups on the oral discourse sub-scores.

8:4 Correlations Between Individual COLT Items and Adjusted Post-test Means

Since the analysis in terms of experiential and analytic groups showed little overall effect of classroom orientation on test performance, the question arises as to whether proficiency differences between classrooms can be related to any of the individual observation variables identified by the COLT. The purpose of this section is to explore the empirical relationships between COLT categories and proficiency outcomes, without any a priori assumptions of what these relationships might be.

We used the classroom as the unit of analysis, since the COLT scheme applies to each classroom as a whole. Thus there were eight observations in the sample, and only quite high correlations are statistically significant. Indeed, the number of correlations reaching significance is not above the chance level (one in twenty) for this number of correlations. However, since the purpose of this section is to explore the data and to suggest possible relationships for future study, it seemed sensible to pay attention to the patterns of correlations that did occur, and to try to interpret them, even though the results fall short of significance.

Before looking at the relation between COLT categories and proficiency, let us look at the relationships among the proficiency variables themselves. As proficiency measures, we take the adjusted post-test mean for each classroom. Table 15 (p. 119) shows the intercorrelations (Pearson) among these measures, and the correlation of each with the total orientation score from the COLT, i.e., the score that was used earlier to classify classrooms as more experiential or more analytic, a higher score indicating that the orientation was more experiential.

First, in agreement with the overall results of the previous two sections, the COLT pedagogic orientation score does not correlate significantly with improvement on any of the proficiency measures. Even the marginal relation with the written grammar score does not hold up when classrooms are considered separately. However, the intercorrelations between adjusted post-test means are in many cases quite high, showing that improvement is not haphazard. There tends to be a split between the four written measures and the four oral measures: the correlations within either set of scores are higher than those across sets of scores. This means that if a classroom shows strong improvement in one written measure, it tends to show improvement in the other written measures as well; and similarly (although less strongly) for the oral measures. But the classrooms that improve most in written measures are different from those that improve most in oral measures.

Part I categories

Mean values and ranges for each of the COLT Part I categories are presented in Table 16 (p. 120), and correlations between each of the Part I categories and the various improvement measures (i.e., adjusted post-test means) are shown in Table 17 (p. 121). In Table 17, a positive correlation indicates that the classrooms that spent relatively more time on the designated COLT activity improved more in the designated proficiency measure or, in other words, that greater than average time devoted to the activity promoted learning of the designated sort. A negative correlation indicates the opposite, that greater than average time devoted to the activity interfered with learning of the designated sort. Perhaps this interference was direct in some cases, but it is probably more likely that the less productive activities took away time from more productive activities. In this regard, it should be remembered that within major headings of COLT Part I, the time spent on the various sub-categories generally adds up to 100%, so that within a major heading one is always looking at the relative effectiveness of different uses of time.

In some cases, the distribution of classrooms on a given Part I activity is such that only one or two classrooms spent any appreciable time on the activity in question. In these cases, the correlation is not worth considering since it is based on only one or two cases, and blanks have been shown in Table 17. In the remaining cases, there is a fairly even distribution of time spent across different classrooms within the range shown in Table 16.

Consider the correlations for COLT Part I section by section, beginning with participant organization. There is not enough time spent on group work (categories 6-7) to draw any conclusions, nor on individual work with students involved in different activities (9). With regard to the other participant organization categories, classrooms in which relatively more time was spent with the teacher addressing the class (3), or with students working individually on the same activity (8) showed relatively good improvement on most proficiency measures, while classrooms which spent more time on student/whole class (4) and choral work (5) showed relatively little improvement. This may mean that these activities do not contribute to language learning. On the other hand, as mentioned above, it may be that student/whole class and choral work are not so much bad in themselves, but rather that they take time away from the more effective types of organization. In these results, the only significant relationships are the negative ones between student/whole class activities and listening comprehension and oral sociolinguistic performance, but the consistency in direction of the other relationships makes them worth noting for future consideration.

With regard to classroom content, there is not enough time spent on language categories other than form (13-15), nor on topics with narrow range of reference (16-19), to produce reliable data. Time spent in classroom management (10-11) is somewhat positively related to improvement in the written measures and grammatical oral. Concentration on formal aspects of language (12) is somewhat positively related to most measures of improvement. Discussion of topics other than classroom management and language (20-29), shows mixed correlations with improvement, although discussion of topics with a limited scope is perhaps detrimental. The categories of content control (30-32) are essentially unrelated to improvement, indicating that who selects the topic that is being talked about — the teacher, the student or both — is not important.

The impact of different student modalities is mixed, except that time spent on student talk (34) is consistently detrimental to improvement, perhaps because in classrooms where an especially large amount of attention is given to this activity the time is not structured well enough to be useful. Type and source of material does not seem to be strongly related to improvement, except that frequent use of visual aids (41) and perhaps L2 materials (42) appears to be useful.

Part II categories

Results for COLT Part II are presented in Tables 18, 19a, and 19b (pp. 122 - 124). The first category in Table 19a — teacher off-task activity — is interesting, especially in comparison to student off-task activity (19). Teacher off-task activity tends to correlate positively with learning, suggesting that the seemingly irrelevant comments of teachers when they are off-task may actually present important learning opportunities. In contrast, time spent in student off-task activities tends to be detrimental to learning.[23]

Use of L1 or L2 by teacher (2-3) or students (20-21) is not related strongly to learning, but it should be noted that in all the classrooms a great percentage of the time is devoted to L2.

In the next group of teacher categories, concerning information gap, it is interesting to note that the narrow gap categories, predictable information (4) and display request (6), are uniformly detrimental to learning. Unpredictable information (5) shows mixed correlations with learning measures, while correlations of genuine information requests by the teacher (7) with oral measures are positive, suggesting that this type of activity is valuable for oral improvement. Sustained speech (9) on the part of the teacher is positively correlated with improvement, especially for discourse oral, suggesting the importance of the teacher as model.

Reaction by the teacher to the linguistic code (10) seems to be relatively important to improvement in oral production measures, while reaction to message (11) seems to be more important to improvement in written production measures, although these correlations are not extremely high.

The next group of categories refers to topic incorporation. Correction of student utterances by the teacher (12) uniformly shows a negative impact on learning. Paraphrases (14), elaboration requests (18) and especially clarification requests (17) by the teacher promote learning. Repetitions (13), comments (15), and expansions (16) are relatively neutral. These results suggest that the important variable for learning is that the teacher help the students to express their own ideas.

Consider now the student categories (Table 19b). Discourse initiation by students (22) is not particularly related to improvement. Predictable information giving (23) and genuine information requesting (26) are both negatively correlated with improvement. It is possible that genuine information requests may signal confusion on the part of students. What the source of confusion might be, of course, cannot be determined from our results.

Ultraminimal student speech (27), which reflects the traditional transmission mode of teaching, tends to be positively correlated with written measures, and negatively with oral. Minimal (28) and sustained (29) student speech are both somewhat negatively correlated with improvement, in agreement with the negative correlations between student speech and improvement in COLT Part I. Sustained student speech is perhaps less negatively correlated with improvement than is minimal student speech, suggesting at least some relative benefit when student utterances are substantial enough to make a contribution to classroom discourse.

Use of restricted forms by students (30) is negatively correlated with learning, especially on written tests. Reaction to message (33) is somewhat positively related to learning, while reaction to code (32) tends to be negative.

All categories of topic incorporation by students (34-40) tend to be correlated with improvement on written measures. However, these categories show mixed correlations with improvement on oral measures.

Combined categories

Table 20 (p. 125) shows correlations between the proficiency measures and a number of combined COLT categories. The score for each combined category was calculated as the simple sum of time spent in all the constituent categories. The combined categories were constructed in pairs, in order to determine in each case which of two opposed processes is the more effective.

The first comparison is between (a) three types of whole class activity and (b) four types of group/individual activity. Here the correlations are generally low with all of the proficiency measures, indicating that this dimension is largely unrelated to improvement. Referring back to the constituent categories, one can see the reason for the lack of correlation between whole class work and improvement. The correlation (Table 17) for whole class with teacher talking is positive, while the correlation for student talking is negative. In other words, the important variable is not whether the whole class is involved in the interaction, but whether they are listening to the teacher or to other students.

The second comparison is between (a) focus on form and (b) message-focused activities with limited or broad range of reference. Focus on form tends to be positively related to improvement, while the discussion of general topics with limited and broad reference tends to be detrimental. As indicated in the analysis of the individual categories, control of content by teacher, by student, or jointly controlled by teacher/student (comparison 3) is not important to improvement.

The next pair of combined categories (comparison 4) suggest that involvement in extended writing is more beneficial to written proficiency than sustained speech, but the opposite does not hold for improvement in oral measures. None of the correlations are very high, however. The source of materials used (comparison 5) is not strongly related to improvement.

Information gap (comparison 6) is related to improvement on written measures, with predictable content/display request on the part of teacher and student negatively correlated with performance, and unpredictable content/information request positively correlated. This pattern, however, does not hold for oral measures. Reaction to code vs. reaction to message (comparison 7) shows a negative relation for reaction to code, and a positive relation for reaction to message, which holds across all proficiency measures.

The final comparison shows that both the relatively restricted and the relatively expanded ways of incorporating utterances are somewhat positively related to improvement on the proficiency measures. More detailed analyses of topic incorporation have already been given under the individual categories, which suggest that the most important variable for teachers is not restricted vs. expanded, but perhaps something more like encouraging vs. discouraging. The generally positive correlations for both the restricted and the expanded categories suggest that any incorporation of utterances, or in other words the very existence of give and take between teachers and students in the classroom, is beneficial to learning.

9. DISCUSSION AND INTERPRETATION OF RESULTS

The results will be discussed in three stages. First, we will interpret the proficiency results with respect to our specific hypotheses regarding the interaction of classroom variables and student performance. At the second stage, we will interpret the correlations between COLT items and adjusted post-test scores in the light of current communicative language teaching theory. At the third stage, we will look more closely at the transcripts of class 2 and class 5, in order to consider the hypothesis that it is the quality rather than the quantity of interaction which aids development.

9:1 Comparison of Experiential and Analytic Groups

The most striking aspect of the results, one which runs counter to our predictions, is the extent to which Groups A/A* and E are indistinguishable. We find statistically significant differences for only a small number of tests, namely the grammar test, and the letter and note. The comparisons here involve the total scores. As well there were significant differences on various subtests within each general test-type. However, on most subtests and on the fluency and strategic competence measures, the groups were basically the same.

Hypothesis 1A, which predicted that Group A (or Group A*) would score higher on the multiple choice grammar test, was confirmed. In interpreting this result it should be recalled that group A* did not spend significantly more time than Group E on explicitly form-focused activities such as oral pattern practice, written sentence conversion exercises, or the use of overt grammatical explanations. Nevertheless, Type A classrooms were significantly different from Type E classrooms with regard to a number of features which are generally associated with the traditional transmission style of teaching. These features included topic selection by the teacher, and various characteristics of student discourse, i.e., minimal written text, minimal utterance in spoken interaction, reaction to code rather than message, and restricted choice of linguistic item. Conversely, Type E classrooms spent significantly more time on activities marked by student topic-selection, extended written text, sustained speech, reaction to message rather than code, topic expansion, and use of student-made materials, which are features generally associated with a reciprocal interaction approach (cf. section 5:2). These results suggest that a number of activities not directly related to form-focused practice may have had a beneficial effect on grammatical proficiency.[24]

We have already indicated that Type A and Type E classrooms differed in their over-all pedagogic orientation. In the light of this difference, the results on the multiple choice grammar test suggest that a relatively strong analytic focus may lead to a certain level of mastery of grammar which is not replicated by a more experiential approach. In a pedagogic context where students lack access to native speakers, where input and opportunities for practice are limited, and input exposure time is also constrained, it appears that the provision of comprehensible input may not be enough to guarantee the acquisition of grammatical generalizations.

A lack of significant differences on the various grammar measures of the oral test runs counter to Hypothesis 1B. The two sets of results therefore suggest that measurements of grammatical competence are dependent on the medium of testing, Group A* doing better when recognition rather than production tests are used.[25] They do not outperform Group E on tests which require them to activate their knowledge on the basis of what they know themselves. This finding calls to mind the distinction made by Bialystok and Sharwood-Smith (1985) between knowledge and control. The first term refers to the way in which the language system is represented in the mind, while the second refers to the processing system which controls the knowledge system during actual performance. In the case of our Group A* students, we may conclude that they have greater grammatical knowledge but not greater control. This conclusion, of course, is compatible with claims that L2 practitioners have made repeatedly over the years: the acquisition of grammatical knowledge does not guarantee the student's ability to put that knowledge to use for purposes of normal communication.

The various subanalyses indicate that Group A* students do not outperform the Group E students on all grammatical measures. The two groups come out about even on oral and written measures of subcategorization and gender. The one area of grammar where Group A* obtained higher scores on the written grammar test was with respect to agreement processes. There was no significant difference on the oral test on agreement. This is interesting in that patterns of agreement in written work represent the type of error which has traditionally received a great deal of pedagogic attention. A greater emphasis on reaction to code (e.g., correcting grammatical errors) in Type A classrooms could explain our results here. In contrast, subcategorization and gender involve properties of the grammar which must be learned along with individual words. Systematic exercises focusing on gender or the subcategorization properties of verbs are not part of traditional grammar teaching,[26] and it may be the case that acquiring this kind of grammatical knowledge is not facilitated by the types of structural exercise which are commonly found in core French classrooms.

Hypotheses 2A and 2B were roundly disconfirmed. There were no significant differences between the groups on the written or oral discourse measures. Furthermore, although the overall results on the written measures were quite good, neither group showed evidence of improvement during the year. We attribute these results to the virtual absence of instruction on discoursal aspects of language in both types of classroom. In the absence of explicit instruction, we cannot assume that knowledge of these aspects of language will be inferred from more general features of experiential teaching. We conclude, therefore, that explicit discourse-related instruction is necessary to ensure that students acquire knowledge of discourse rules.

Hypotheses 3A and 3B were also disconfirmed. There were no significant differences on the contrastive aspects of the sociolinguistic features we examined, although Group A* did show greater mastery of the specific formal markers necessary to establish the contrast between a formal letter and an informal note. Neither group showed much evidence of improvement during the year on written or oral sociolinguistic measures. Again, we attribute these results to the virtual absence of explicit sociolinguistic instruction in both types of classroom, with the same general conclusions.

Hypotheses 4A and 4B were disconfirmed. There were no significant differences between the two groups on the listening comprehension test, nor did the groups differ in their ability to understand questions during the oral interview. There are two possible explanations for these results. The first point to note is that students in Type A and Type E classrooms spent more time listening (45.82% and 52.29% respectively) than on any other type of modality. This may explain the students' favourable assessment of their own comprehension skills as opposed to their speaking and writing skills, and it may account for the lack of significant differences between the two groups on the listening comprehension measures. A second possibility is that neither group spent much time listening to authentic materials, particularly those marked by a variety of accents, styles of delivery, etc. If it is the case that there was very little use of naturalistic speech other than that produced by the teacher in either type of classroom, this could help to account for the lack of differentiation between the two groups.

Hypothesis 5 stated that students from experiential classes would show higher accuracy rates on the oral sentence repetition task. This hypothesis was disconfirmed, a result which is consistent with the results of the listening comprehension test. Neither Type A* nor Type E students showed much ability to parse (i.e., segment speech) when they were listening to naturalistic language, or to language which had at least the phonological features of naturalistic text.[27] We attribute this result to the general lack of opportunity afforded either group to listen to sustained speech marked by normal features of rhythm and prosody. In view of the difficulties most students experienced during the oral interview, we suggest that this property of the classroom environment (i.e., the availability or otherwise of naturalistic listening practice) should constitute a major focus for future research.

Given our present sample of core French students, it does not appear that an overall focus on experiential activities constitutes a necessary and sufficient condition for the development of communicative skills. In the case of our Group E students, a relatively strong experiential focus did not result in superior communicative skills, since Group E were not significantly different from Group A* on the discourse, sociolinguistic, and oral fluency measures. Furthermore, a relatively strong experiential focus did not appear to produce comparable grammatical proficiency, since Group A* was significantly different from Group E on the multiple choice grammar test. Evidently, it is possible for some students to achieve an equal level of communicative skills by participating in a program which has a relatively strong analytic focus. On the other hand, it cannot be assumed that students will automatically develop equal or superior grammatical competence by spending relatively more time on experiential, message-oriented activities. These conclusions must be tentative, since none of the classrooms in our sample were located at the extreme ends of a prototypic high-low pedagogic scale. It could be argued, however, that our sample represents the average conditions under which a great deal of core French instruction takes place in Ontario. At the very least, our study raises the question of whether we are wise to assume — as many writers currently appear to do — that the more innovative aspects of communicative language teaching can be applied in general terms to all types of clasroom, regardless of such factors as the personality of the teacher, the needs of the students, or the amount of time available for instruction.

9:2 Correlations Between Individual COLT Items and Improvement Scores

In our comparison of COLT categories and performance measures we looked first at the relationship between adjusted post-test scores and individual COLT categories, and then at the correlations between adjusted post-test scores and various combinations of categories. Taking the individual categories first, we found that the profile of a successful classroom which emerged from COLT Parts I and II was as follows: the teacher does relatively more talking compared with individual students to the class as a whole; relatively more time is spent on classroom management; more time is spent on form-focused activities than on general discussion; the students themselves spend relatively little time speaking; and visual aids and L2 materials are used relatively often. The analysis based on combined COLT categories showed that focus on form, extended writing, information gap, reaction to message, and topic incorporation were positively related to improvement, while sustained speech by students, predictable content/display request, reaction to code, and general discussion with limited or broad range of reference were negatively related. In the case of participant organization, topic control by teacher or student, and use of 'authentic' L1 or L1-adapted materials, the correlations with performance measures were generally low, indicating that these aspects of classroom treatment were relatively neutral with regard to improvement. As already mentioned — and we want to emphasize this fact — few of the correlations were statistically significant. However, our purpose was to look at all the patterns that seemed to show consistency, in order to identify possible relationships for future study.

The above classroom profile indicates that our core French students benefited from a generally experiential approach in which relatively more time was devoted to such features as information gap, reaction to message, and topic incorporation. At the same time, there were positive correlations between various form-focused, teacher-directed activities and adjusted post-test scores. It is possible to interpret these results as lending support to a 'weaker' or more conservative version of communicative language teaching, according to which experiential activities serve as an enrichment of a basically form-focused program, rather than to a 'stronger' or more radical version, according to which there is no need to provide systematically graded input, since it is assumed that grammatical knowledge will develop automatically out of spontaneous language use (cf. Johnson 1982).

As we have indicated, our results must be interpreted with caution. Clearly, before we can hope to draw general conclusions it will be necessary to replicate the study in many different contexts, using either the existing COLT variables, or other combinations of pedagogic features. In the meantime, however, our results serve as a useful reminder that the patterns of classroom interaction are extremely complex, and that this complexity is difficult to reconcile with the sweeping generalizations, often unrelated to any specific instructional setting, which are frequently found in the current language teaching literature. The implications for curriculum development, tentatively stated, are that the analytical focus and the experiential focus may be complementary rather than two ends of a continuum, and that they may provide essential support for one another in the L2 classroom.[28] Future research should focus on this issue. In particular, we would like to see similar studies conducted in instructional settings where significant differences may be found on a wider range of variables than those identified in the core French program.

As previously indicated (section 5:2), those aspects of COLT which proved most useful for distinguishing between the classrooms in our sample were topic selection, extended and minimal written text, reaction to message/code, sustained and minimal speech, form restriction, source/purpose of materials, and topic incorporation. Significant differences were not found with respect to focus on form or other aspects of language, range of reference, participant organisation, student modality, or information gap. Furthermore, there were several types of activity which occurred either not at all or very rarely during observed time in both types of classroom. These infrequently observed activities included group work, choral work, student discourse initiation, and explicit focus on function, discourse or sociolinguistics. It is possible that a process-product study conducted with a sample of classrooms from a more 'communicative' program — one where students spend more time talking on general topics, where teachers provide explicit focus on discoursal and sociolinguistic aspects of language, and where there is more group work alternating with teacher-directed whole-class activities — would provide us with a better chance of identifying the treatment factors which are most relevant to particular aspects of learning.[29]

9:3 Quality vs. Quantity of Interaction

Finally, we calculated the total gain in proficiency for each school over the year (see Table 21, page 126). Of the two experiential classes one (class 2) made the highest gain in overall proficiency and the other (class 5) made the lowest gain (cf. Table 6, p. 103). What characteristics of these two classes were responsible for the striking difference in proficiency results? Ellis (1984) suggests that it is not the quantity of interaction that counts but the quality, and formulates two hypotheses: (a) development is fostered by consistency and accuracy of teacher feedback; (b) communicatively rich interaction which affords opportunities for the negotiation of meaning may aid development, where more structured forms of interaction do not. In order to examine these ideas, we undertook a qualitative analysis of the transcripts for classes 2 and 5.

The qualitative analysis provided evidence that the high-scoring experiential class engaged frequently in communicatively rich interaction, involving feedback and the negotiation of meaning. The low-scoring experiential class, on the other hand, received less feedback and spent more time on stereotyped routines which lacked the quality of spontaneous discourse. The difference between meaningful interaction and stereotyped routines can be illustrated with reference to the teaching of grammar. Class 2 spent 65.64% of observed time on activities which involved a focus on formal features of language, but this was usually done in the context of meaningful tasks such as correcting the errors in student composition. Such tasks, directed by the teacher at the blackboard, provided an opportunity for everyone to work together on developing ideas and finding the best way to express them in the target language:

(1) T: comment est-ce que vous dites 'Mr. Reagan does not sleep too much?' (...) okay quel temps est-ce? pensez (name)
S: conditionnel? (...)
T: oui conditionnel de quel verbe? (name)
S: dormir

T: donc, M. Reagan ne ... continuez
S: devrait pas
T: devrait pas? comme ça? est-ce correct?
S: non
T: NON! quel est l'idée de l'anglais? (...) et ici c'est le problème toujours dans les compositions quand vous exprimez toujours quel est l'idée de dans l'original? (...) Okay expliquez-moi ça en anglais c'est une expression anglaise vous savez pas est-ce qui oui ça c'est que j'ai dit et vous savez pas votre anglais ... voilà (name)
S: um um juste je dis que la forme est ah il y a la
T: oui
S: uh comme M. Reagan il faut que M. Reagan dort dorme ne dormera pas tard
T: oui c'est une moitié vous avez l'idée est bonne l'idée est bonne c'est l'expression de l'idée est un peu faible maintenant...
(cl.2/obs.1/pg.20)

Class 5 spent 23.89% of observed time on form-focused activities, but in this case the activities often consisted of 'decontextualized' grammar practice which was clearly lacking in genuine communicative intent:

(2) T: okay? une question avec qui est-ce qui (name) on peut employer qui ou qui est-ce qui sujet du verbe
S: qui
T: qui est-ce qui
S: qui est-ce qui
T: qui est-ce qui est à la porte je je n'sais pas est-ce qu'il y a quelqu'un?
S: non
T: non qui est-ce qui uh qui est-ce qui est ton professeur okay (name) qui est-ce qui est ton professeur de mathématique?
S: mon professeur de mathématique est M. (name) (...)
T: M. (name) est mon professeur on peut dire qui est mon professeur ou bien qui est-ce qui ... une question avec comment (name) ... pose-moi une question avec comment
S: um (laughter) uh comment um
T: comment
S: (laughter) comment ça va?
T: comment ça va ça va très bien uh comme-ci comme-ça ...
(cl.5/obs.1/pg.7)

The drill-like practice illustrated above contrasts with the technique used by the class 2 teacher. According to the 'quality interaction' hypothesis, the class 2 procedure is likely to be pedagogically more effective since it emphasizes meaning negotiation and the development of metalinguistic awareness.

In the case of class 2, the most striking examples of jointly-negotiated meaning occurred in a lesson devoted to a philosophical discussion of Le Petit Prince. In this discussion the teacher insisted that the students use French to develop and express their own ideas, thus helping them to establish links between the text and the world of their own experience:

(3) T: bonne question pourquoi est-ce qu'il est choqué?

S: parce qu'il sait que le corde est inutile parce que dans sa planète uh tout est trop petit
T: no! no! no! vous avez raison ce que vous dites est correct MAIS ce n'est pas la raison qu'il est choqué
S: parce que le petit prince est très gentil
T: oui! continuez il est très gentil
S: um il um aime le mouton
T il aime le mouton okay continuez ça c'est bien
S well freedom
T: ah voilà! okay comment est-ce qu'on dit 'freedom'?
S la liberté
T: okay parlez-moi un peu de la liberté et mouton (laughter) oui faites cette connection oui?
S: le il veut la liberté pour le mouton
T okay il veut la liberté pour le mouton seulement pour le mouton?
S: non pour tout le monde
T: ah voilà eh! nous touchons maintenant à un autre aspect un autre caractéristique du petit prince ...
(cl.2/obs.3/pp. 78—80)

The communicatively rich interaction which resulted from discussing Le Petit Prince in the high-scoring class contrasted with the stereotyped nature of student presentations in the low-scoring class. In one class 5 lesson several groups of students gave presentations on topics of general interest such as 'videos', 'abortion', and 'popular TV programs'. The students had prepared the topics themselves without help from the teacher, and the activity was potentially a valuable one. Unfortunately, however, the students addressing the class articulated so badly that it was difficult to hear what they were saying, the discussion which followed each presentation was generally in English rather than French, and the teacher provided virtually no feedback concerning the students' use of the target language. A review of the transcript makes it clear that these factors must have seriously detracted from the effectiveness of the activity. Nevertheless, in terms of the COLT coding scheme, the classroom presentations received credit for such experiential features as extended speech, broad range of reference, and content control by the students.

It appears, then, that a statistical analysis based on COLT cannot be depended on to distinguish between pedagogically effective communicative activities, and pedagogically ineffective routines which may bear only a superficial resemblance to normal conversational behaviour. One problem is that the COLT observation scheme was designed to provide a broad picture of the types of activity which characterize L2 classrooms. As a result, it does not enable us to pay sufficiently close attention to the exchange structure of discourse, particularly to the way in which conversations are jointly negotiated by means of various topic incorporation devices. The importance of topic incorporation in facilitating mother-child interaction has been clearly demonstrated (Wells 1985), and there is reason to believe that it plays an equally important role in second language acquisition (Pica 1987). In any future study, therefore, it is important that the observation procedures based on COLT be supplemented by a more detailed discourse analysis, with a view to obtaining additional information about the way meaning is negotiated in the classroom.

Footnotes

1. Stern (1978) distinguishes between learning a language through use in the environment (i.e., functionally), or through processes of language study and practice (i.e., formally). As Stern points out, this aspect of language behaviour can be characteristed as a psycholinguistic/pedagogic continuum, or 'P-scale'. There is nothing inherently good or bad about activities at either end of the scale, and in organized language teaching we often find an interplay between formal and functional approaches. In this study the term 'experiential' is used to refer to activities at the functional end of Stern's P-scale, while 'analytic' refers to activities at the formal end. The experiential-analytic distinction is analogous (although not necessarily identical) to distinctions made by other investigators with regard to general pedagogic orientation. Barnes (1976), for example, discusses 'interpretive' versus 'transmission' teaching; Wells (1972) distinguishes between 'collaborative' and 'transmission' orientations; while Cummins (1984) labels these dimensions 'reciprocal interaction' versus 'transmission'.

2. The communicative language teaching literature suggests a simple dichotomy between types of classes, i.e., a structurally-oriented, teacher-controlled class cannot be simultaneously communicative, and a functionally-oriented, student-controlled class cannot be simultaneously analytic. One of our objectives was to investigate the well-foundedness of this viewpoint. It should be emphasized, however, that the COLT categories are binary and not scalar so that with respect to any one feature, a given class could be only experiential or only analytic. Nevertheless, a class might be judged more-or-less communicative through various combinations of experiential features and analytic features. We would argue that one of the advantages of the COLT is precisely that it can make these finer-grained distinctions.

3. When the entire sample was considered, the high-contact analytic group scored higher ($p < .05$) than the high-contact experiential group on the multiple-choice grammar test, and the low-contact analytic group scored higher than the low-contact experiential group ($p < .05$) on the same test. When the interviewed subsample was considered, the high-contact analytic students scored higher than the low-contact analytic students ($p < .01$) on the multiple-choice listening test. This result should be interpreted with caution, however, since it was based on only 16 students, none of whom were characterized as high-contact experiential.

4. Simultaneous testing was necessitated by the fact that all testing had to be done during the French period. We were requested not to keep students from other classes.

5. Given the location of the study, it would have been pointless for the testers to pretend not to know English. In any case, during the interview we specifically wanted to see which students, if any, would resort to English as a communicative strategy.

6. Despite all of our assurances to the contrary. As one student put it "Ce n'est pas un test pour toi mais c'est un test pour moi!"

7. The enrichments involved phonetic transcriptions of anglicisms, borrowings, partially nativized words and phrases, and other phonetic curiosities. Also recorded were pauses, interruptions or self-corrections, focus, missing or deviant liaison, and failure to reduce vowels in articles or pronouns occurring before vowel-initial words (e.g. (le # article, je # aime). We were not interested in establishing that the students did or did not have an accent when they spoke. Phonetic (as opposed to phonemic) deviations were ignored in transcribing.

8. The coding procedure consisted of the identification of all non-pronominal noun phrases which were bracketed using pairs of 'NP' symbols, e.g. J'ai vu NP-ma copine-NP. Subject-verb sequences were bracketed using pairs of 'V' symbols, e.g. V-NP-le facteur-NP va-V dans NP-la rue-NP. In the latter case, only those instances where the third person singular and plural forms of the verb were dissimilar were singled out since we wanted to examine agreement phenomena. Auxiliaries were not counted.

9. For a number of reasons it was not possible to calculate intercoder reliability coefficients. In any future study, however, we recommend that intercoder agreement should be determined statistically.

10. The grouping of categories into experiential and analytic was based on a review of the communicative language teaching literature (see Chapter 3). It is important to emphasize that the pedagogic orientation of classrooms is not determined by a single feature, but by a cluster of interrelated dimensions. For example, it would not make sense to take the single feature 'group activity vs. whole-class activity' and to use it as the basis for distinguishing between experiential and analytic classrooms. However, if we find classes where relatively more time is spent on a combination of activities marked by group work, broad range of reference, use of extended text, reaction to message rather than code, etc., it is possible to characterize these as having an over-all experiential profile. Similarly, classrooms which spend relatively more time on whole-class activities, form-focused practice, use of minimal text, reaction to code rather than message, etc., can be described as having an over-all analytic profile.

11. Class 5 was a large, ethnically mixed class which presented a number of discipline problems.

12. Our assumption is simplistic since text can be in principle both extended, i.e. more than a syntactic phrase, and linguistically simplified. In the absence, however, of any independent measures of the naturalness of the French heard in the class we contented ourselves with this relationship since even relatively slow discourse will manifest some of the features mentioned above.

13. Listening comprehension presumably involves at least three types of competence and we presuppose the existence of strategies for decoding text too.

14. In some cases, one of the choices involved selecting no word at all.

15. The focus of the grammar test was on obviously learned and learnable (hence teachable) aspects of syntax and inflectional morphology. These aspects contrast

with universal (and potentially innate) features of grammar. Since we were interested in the effects of environment on learning, we were not interested in establishing the mastery of those aspects of grammatical competence which depend on universal principles such as the c-command constraint on anaphoric binding, or the subjacency condition on filler-gap dependencies. Statements about the grammatical competence of the students should be interpreted with the awareness that we are limiting our discussion to only a small part of the linguistic system.

16. One can argue that once a learner has realized that words ending in -ation bear a suffix that gets familiar gender, then the learner should generalize feminine gender to all instances of that suffix, even to words that the learner has never heard before. In the case of morphologically unanalysable words, however, there are no structural reasons which would lead the learner to select a particular gender. This does not mean, of course, that learners will not exhibit strategies of gender assignment when dealing with unknown words, a topic that deserves further study.

17. This decision was less arbitrary than might first appear. Clearly two features were necessary for there to be any ordering. It was felt that the request and the rationale could appear in any order relative to each other. Consequently, they could only be fixed with respect to the opening and closing, a fact which meant that we were looking for only a minimal amount of internal ordering.

18. The particular choice of items was selected in order to permit future comparisons with studies of other populations, see in particular Harley 1986.

19. Scoring for this feature was fairly straightforward since the interviewer supplied a verbally unsolicited repetition of the question. The interviewer's interpretation of the need for the repetition may, however, have been incorrect.

20. The symbol ‿ encodes liaison heard on the stimulus recording.

21. In discussing these data we shall rely heavily on adjusted post-test scores as measures of the relative success of the two groups. These scores should be understood as the post-test score that each group would have achieved, had the two groups been equal at pre-test.

22. Separate tables for Groups E and A* are not provided, since there were few differences between the two sets of results.

23. This result makes sense in view of the fact that a wide range of topics were coded as off-task. Thus, if basketball was referred to in a textbook exercise, and the teacher started talking about the school basketball program, this was coded as off-task because the conversation interrupted the exercise. Similarly, a disciplinary episode which developed into a discussion about the rights and wrongs of gum-chewing was coded as off-task because it interrrupted the activity that the class was currently engaged in. Teacher off-task activity often developed into meaningful interactions, whereas student off-task activity was less usefully structured.

24. The implication is that the relationship between classroom environment and proficiency may be subtle and indirect, and therefore difficult to capture in an observational study.

25. Note that the multiple-choice test was not strictly speaking a 'pure' measure of grammatical competence since students were required to select the proper response from among a set of distractors.

26. The exercises we have in mind would develop, for example, the systematic relationship between l'eau bout and Jean fait bouillir l'eau, and between J'ai décidé de partir, Je suis décidé à le voir, and Il a été décidé de partir.

27. This difficulty was compounded for the students by their relative lack of skill in producing fluent, automatized speech.

28. Our conclusion that analytically-focused and experientially-focused teaching approaches may be complementary in core French programs can be compared with similar findings with respect to French immersion. Thus Harley and Swain (1984) claim that the simple provision of comprehensible input in a classroom setting is not sufficient to ensure productive use of formal aspects of L2, even in situations where students are exposed to the target language for several hours a day, and thus also have more occasions for using the language. Harley and Swain found that while early immersion students do extremely well on tasks involving global comprehension of discourse in the target language by grades 5 and 6 (and this after six or seven years of immersion), they are still making grammatical errors which clearly set their production apart from that of native speakers of the same age. Considerations such as this have led to the suggestion that learning in an immersion setting may be facilitated by providing language input that is "explicitly designed to clarify the meaningful use of particular grammatical forms, and by devising communicative contexts in which students practise the productive use of such forms" (Chapter 5, Appendix A).

29. Any study conducted with other types of program would have to deal with the problem of in-class vs. out-of-class exposure to the target language. Thus, ESL in an English-speaking environment such as Toronto would not be a suitable program for replicating our study, and FSL in Montreal would present the same difficulties.

References

Barnes, D. 1976. **From Communication to Curriculum.** Harmondsworth: Penguin.

Bialystok, E. and Sharwood Smith, M. 1985. Interlanguage is not a state of mind: an evaluation of the construct for second language acquisition. **Applied Linguistics** 6/2, 101-17.

Carrell, P. 1982. Cohesion is not coherence. Paper presented at the Second Language Research Forum, Los Angeles.

Carroll, J.B. 1975. **The Teaching of French as a Foreign Language in Eight Countries.** International Studies in Evaluation V. Stockholm: Almqvist and Wiksell.

Cummins, J. 1984. **Bilingualism and Special Education: Issues in Assessment and Pedagogy**. Clevedon, U.K.: Multilingual Matters.

Ellis, R. 1984. Can syntax be taught? A study of the effects of formal instruction on the acquisition of WH questions by children. **Applied Linguistics** 5/2, 138-155.

Fröhlich, M., Spada, N. and Allen, P. 1985. Differences in the communicative orientation of L2 classrooms. **TESOL Quarterly** 19/1, 27-52.

Halliday, M.A.K. and Hasan, R. 1976. **Cohesion in English.** London: Longman.

Harley, B. 1983. Transfer in the written compositions of French immersion students. **Development of Bilingual Proficiency: Year 2 Report,** Appendix C.

Harley, B. and Swain, M. 1984. The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper and A.P.R. Howatt (eds.) **Interlanguage.** Edinburgh: Edinburgh University Press.

Krashen, S.D. 1982. **Principles and Practice in Second Language Acquisition.** Oxford: Pergamon.

Pica, T. 1987. Second language acquisition, social interaction, and the classroom. **Applied Linguistics** 8/1, 3-21.

Stern, H.H. 1978. The formal-functional distinction in language pedagogy: a conceptual clarification. In. J.G. Savard and L. Laforge (eds.) **Actes du 5e. Congrès de L'Association Internationale de Linguistique Appliquée.** Montreal: University of Laval Press.

Wells, G. 1982. **Language, Learning and Education.** Centre for the Study of Language and Communication, University of Bristol.

Wells, G. 1985. **Language Development in the Pre-School Years.** Cambridge: Cambridge University Press.

Table 1

Length of class and distribution per board*

BOARD LABEL	CLASS** LABEL	TIME IN MINUTES x NO. of CLASSES = TOTAL IN MINUTES PER WEEK
A = 2 classes	Class 1	70 x 3 = 210
	Class 2	70 x 3 = 210
B = 3 classes	Class 3	70 x 3 = 210
	Class 4	70 x 3 = 210
	Class 5	70 x 3 = 210
C = 3 classes	Class 6	40 x 5 = 200
	Class 7	40 x 5 = 200
	Class 8	40 x 5 = 200

* All boards were located in Metropolitan Toronto.

** The class labels are arbitrary.

Table 2

Class size and average age of students

CLASSES	SIZE*	AVERAGE AGE**
1	33	16.9
2	25	16.9
3	23	17.0
4	24	16.9
5	35	16.9
6	30	16.9
7	10	16.8
8	20	16.8

* Class size varied over the course of the year. These figures represent enrolments in May, and come from the teacher questionnaires.

** I.e., average age at the time of the post-tests.

Table 3

Home Language Background of Student as a Percentage Per Class

Class	Patois	Italian	Portuguese	Spanish	Greek	Macedonian	German	Dutch	Arabic	Yiddish	Persian	Hindi/ Urdu	Korean	Chinese	English	Yugoslavian
1			4.0					4.0	4.0	4.0				4.0	80.0	
2		11.8		11.8			5.9						11.8	5.9	52.9	
3					26.7										73.3	
4	5.0					5.0	5.0					5.0			80.0	
5		15.4			19.2	11.5		3.8				3.8	3.8	3.8	38.5	
6		20.0	25.0		10.0								5.0	20.0	15.0	.5
7						11.1									88.9	
8					5.6	5.6					5.6		11.1	5.6	66.7	

Table 4
Contact with French outside the classroom
(expressed as a percentage of total respondents)

1. Have you ever been enrolled in a French Immersion Program? Yes 7 No 93 N.A. 0*

2. Have you ever been enrolled in an Extended French Program? Yes 4 No 94 N.A. 2

3. Please indicate how well you feel you know French by circling the appropriate number for each line on the chart below.

	not at all	with some difficulty	fairly well	quite well	extremely well	N.A.
I understand French	0	23	49	24	3	1
I speak French	1	38	49	8	2	2
I read French	1	24	41	29	4	1
I write French	2	40	42	15	0	1

4. Since the beginning of this school year, how often have you done each of the following in addition to your school work and assigned homework?

		never	hardly ever	some-times	frequently	daily	N.A.
a.	spoken French with your friends outside of school	26	37	30	5	1	1
b.	spoken French with native speakers of French	52	24	16	6	2	0
c.	listened to French music	51	33	14	2	0	0
d.	listened to spoken French on the radio	46	31	19	2	1	1
e.	watched French on T.V.	22	33	37	8	1	1
f.	been to see French movies	74	18	6	1	0	1

* N.A. = No Answer

		never	hardly ever	some- times	frequently	daily	N.A.
g.	read French newspapers or magazines	40	34	22	3	0	1
h.	read brochures or pamphlets in French	34	31	30	4	0	1
i.	read French advertisements or labels on packages	9	13	46	28	3	1
j.	read French books	48	26	20	4	0	2
k.	written letters or notes in French	49	22	20	7	1	1
l.	written other texts in French, e.g. stories	70	23	4	1	1	1

5. Are you presently taking any other French courses in addition to the French classes at school? Yes 1 No 98 N.A. 1

6. Are you presently involved in any extracurricular activities in French, e.g. a French club? Yes 1 No 98 N.A. 1

7. In general, how often have you had the opportunity to use French in some way outside of school during the past five years?

never	hardly ever	sometimes	often	extremely often
12	43	37	5	3

8. How much time have you spent altogether in a French-speaking country (or province) in the past five years?

none	up to 2 weeks	2 1/2 to 4 weeks	4 1/2 to 6 weeks	more than 6 weeks
37	35	10	5	13

9. In the past five years, have you participated in any French activities outside of school, e.g., summer camps?

yes 22 no 77 N.A. 1

10. How much time was involved in total?

up to 2 weeks	2 1/2 to 4 weeks	4 1/2 to 6 weeks	more than 6 weeks	N.A.
13	5	1	3	78

Table 5

Rank order of schools (experiential to analytic) based on COLT Parts I and II

SCHOOL	SCORE	
5	534	Experiential
2	509	
3	400	
7	390	
6	383	
1	358	
8	356	
4	309	Analytic
MEAN	408	

Table 6

Comparative rank order of schools based on COLT score, proportion of time spent on experiential activities, and number of speech turns by teacher and student

COLT score		Proportion of time		Number of speech turns	
School	score	school	time	school	turns
5	534	5	297	5	289
2	509	2	281	2	278
3	400	3	193	3	188
7	390	7	178	7	176
6	383	6	176	6	174
1	358	8	175	8	169
8	356	1	146	1	139
4	309	4	113	4	112

TABLE 7A

Colt part I: Participant organization
Mean percentage of observed time for each category

SCHOOL	OBSV TIME	TYPE	TEACHER SC	STUDENT SC	CHORAL	GROUP SAME	GROUP DIFF	INDIV SAME	INDIV DIFF	GRP INDIV
5	203	E	67.48	30.04	0.00	0.00	0.00	2.46	0.00	0.00
2	262	E	48.09	14.12	0.00	0.00	0.00	37.78	0.00	0.00
	465	E	56.55	21.07	0.00	0.00	0.00	22.36	0.00	0.00
3	299	A	36.95	27.92	0.66	4.34	14.04	15.38	0.66	0.00
7	152	A	61.51	27.30	0.00	0.00	0.00	11.18	0.00	0.00
6	176	A	40.05	43.46	2.27	0.00	0.00	13.06	1.13	0.00
1	275	A	67.45	21.27	0.00	0.00	0.00	11.27	0.00	0.00
8	192	A	59.37	23.43	0.00	0.00	0.00	17.18	0.00	0.00
4	210	A	25.95	28.57	2.85	0.00	0.00	24.04	18.57	0.00
	1304	A	48.19	27.99	0.92	0.99	3.22	15.37	3.29	0.00
	1769		50.39	26.17	0.67	0.73	2.37	17.21	2.43	0.00

TABLE 7B

Colt part I: Content
Mean percentage of observed time for each category

SCHOOL	OBSV TIME	TYPE	MANAGEMENT	FORM	FUNCTION	DISCOURSE	SOCIO LING	NARROW	LIMITED	BROAD
5	203	E	12.80	23.89	0.00	2.70	0.00	0.00	23.15	37.43
2	262	E	6.48	65.64	0.00	0.00	0.00	0.76	4.58	22.51
	465	E	9.24	47.41	0.00	1.18	0.00	0.43	12.68	29.03
3	299	A	6.68	51.33	0.00	0.00	0.00	0.33	24.91	16.72
7	152	A	11.18	57.23	0.00	1.97	0.00	0.00	8.55	21.05
6	176	A	5.68	37.50	0.00	0.00	0.00	0.00	26.98	29.82
1	275	A	7.63	70.06	0.00	0.00	0.97	0.00	12.42	8.90
8	192	A	14.06	48.69	0.00	0.00	0.00	0.00	16.92	20.31
4	210	A	6.19	71.42	0.00	0.00	0.00	0.00	3.33	3.80
	1304	A	8.28	56.95	0.00	0.23	0.20	0.07	16.00	15.79
	1769		8.53	54.44	0.00	0.48	0.15	0.16	15.13	19.27

TABLE 7C

Colt part I: Content control
Mean percentage of observed time for each category

SCHOOL	OBSV TIME	TYPE	TEACHER	TEACHER STUDENT	STUDENT
5	203	E	71.92	3.94	24.13
2	262	E	44.65	38.16	17.17
	465	E	56.55	23.22	20.21
3	299	A	77.59	22.40	0.00
7	152	A	100.00	0.00	0.00
6	176	A	100.00	0.00	0.00
1	275	A	100.00	0.00	0.00
8	192	A	100.00	0.00	0.00
4	210	A	78.09	21.90	0.00
	1304	A	91.33	8.66	0.00
	1769		82.19	12.49	5.31

TABLE 7D

Colt part I: Student modality
Mean percentage of observed time for each category

SCHOOL	OBSV TIME	TYPE	LISTENING	SPEAKING	READING	WRITING	OTHER
5	203	E	65.84	18.31	15.84	0.00	0.00
2	262	E	41.79	11.83	19.46	26.90	0.00
	465	E	52.29	14.66	17.88	15.16	0.00
3	299	A	37.91	23.07	17.34	21.66	0.00
7	152	A	57.40	19.24	16.28	7.07	0.00
6	176	A	44.32	21.21	23.29	7.76	3.40
1	275	A	48.24	12.06	19.81	19.87	0.00
8	192	A	42.31	16.96	22.52	18.18	0.00
4	210	A	35.71	23.73	21.66	18.88	0.00
	1304	A	43.52	19.26	20.00	16.75	0.46
	1769		45.82	18.05	19.44	16.33	0.33

TABLE 7E

Colt part I: Type of materials
Mean percentage of observed time for each category

SCHOOL	OBSV TIME	TYPE	MINIMAL	EXTENDED	AUDIO	VISUAL
5	203	E	4.92	68.96	0.49	0.00
2	262	E	20.99	66.41	0.00	3.81
	465	E	13.97	67.52	0.21	2.15
3	299	A	53.67	27.75	1.67	9.19
7	152	A	34.53	41.44	0.00	10.85
6	176	A	52.84	25.56	0.00	3.40
1	275	A	47.45	34.90	0.00	9.27
8	192	A	48.69	27.86	0.00	11.97
4	210	A	52.85	5.00	5.00	5.71
	1304	A	49.15	26.91	1.18	8.47
	1769		39.90	37.59	0.93	6.81

TABLE 7F

Colt part I: Source of materials
Mean percentage of observed time for each category

SCHOOL	OBSV TIME	TYPE	L2	L1	L1ADAPT	STUMADE
5	203	E	33.99	0.00	0.00	24.13
2	262	E	42.36	0.00	0.00	36.64
	465	E	38.70	0.00	0.00	31.18
3	299	A	49.16	1.33	0.00	18.39
7	152	A	53.94	7.89	0.00	0.00
6	176	A	62.50	0.00	9.65	0.00
1	275	A	68.36	0.00	0.00	0.00
8	192	A	53.12	0.00	8.85	6.25
4	210	A	37.61	0.00	0.00	3.33
	1304	A	54.29	1.22	2.60	5.67
	1769		50.19	0.90	1.92	12.37

TABLE 8A

Colt part II (teacher talk): Use of target language
Mean percentage of observed time for each category

SCH	TYPE	OFF TASK	L1	L2
5	E	2.46	5.36	38.16
2	E	2.69	0.14	60.04
	E	2.58	2.58	49.81
3	A	2.81	3.74	59.82
7	A	4.60	0.30	59.50
6	A	1.70	0.50	60.35
1	A	4.00	2.83	54.82
8	A	2.41	4.64	49.59
4	A	3.25	0.00	46.86
	A	3.10	2.16	54.62
		2.95	2.29	53.19

TABLE 8B

Colt part II (teacher talk): Information gap
Mean percentage of observed time for each category

SCH	TYPE	PREDICT	UNPREDIC	DISPL REQ	INFO REQ
5	E	1.24	9.28	3.17	7.91
2	E	0.34	14.49	2.86	4.33
	E	0.76	12.06	3.00	6.00
3	A	4.36	18.67	2.37	9.27
7	A	0.52	14.28	0.00	8.78
6	A	1.51	9.85	2.55	3.01
1	A	0.00	4.95	2.17	9.62
8	A	0.79	11.34	0.59	5.18
4	A	3.73	10.64	4.25	4.60
	A	1.77	11.21	2.05	6.73
		1.47	11.46	2.33	6.51

TABLE 8C

Colt part II (teacher talk): Sustained Speech
Mean percentage of observed time for each category

SCH	TYPE	MINIMAL	SUSTAINED
5	E	16.10	17.41
2	E	22.93	34.05
	E	19.74	26.27
3	A	36.41	23.22
7	A	25.78	33.10
6	A	35.95	24.39
1	A	22.74	31.55
8	A	29.36	19.09
4	A	24.81	19.86
	A	28.89	24.92
		26.17	25.32

TABLE 8D

Colt part II (teacher talk): Reaction to message/code
Mean percentage of observed time for each category

SCH	TYPE	CODE	MESSAGE
5	E	0.51	4.83
2	E	1.92	3.99
	E	1.26	4.38
3	A	3.57	0.90
7	A	3.48	2.99
6	A	4.22	3.16
1	A	6.58	3.92
8	A	3.68	3.07
4	A	1.35	0.59
	A	3.90	2.47
		3.12	3.04

TABLE 8E

Colt part II (teacher talk): Incorporation of preceding utterances
Mean percentage of observed time for each category

SCH	TYPE	CORRECTION	REPETITION	PARAPHRASE	COMMENT	EXPANSION	CLARIF REQ	ELABOR REQ	UNINTEL
5	E	2.80	5.86	0.75	4.63	4.00	2.30	1.79	4.65
2	E	2.41	3.50	5.08	5.84	3.95	6.36	8.60	3.05
	E	2.59	4.65	3.05	5.28	3.97	4.47	5.42	3.79
3	A	5.76	4.56	0.49	12.57	3.79	1.77	1.48	0.18
7	A	3.51	6.67	5.03	6.37	9.93	3.01	2.76	0.61
6	A	3.66	3.91	1.72	13.18	7.07	2.80	9.18	0.00
1	A	2.00	4.40	1.74	8.59	3.39	3.47	6.16	0.52
8	A	1.98	6.43	0.46	8.21	2.21	4.28	4.09	1.13
4	A	7.42	3.08	0.00	7.66	1.69	0.84	2.37	2.18
	A	3.97	4.79	1.40	9.40	4.30	2.74	4.41	0.80
		3.56	4.75	1.89	8.17	4.20	3.25	4.71	1.69

TABLE 8F

Colt part II (student talk): Use of target language
Mean percentage of observed time for each category

SCH	TYPE	OFF TASK	L1	L2
5	E	0.71	6.14	59.96
2	E	0.00	1.62	37.81
	E	0.33	3.73	48.16
3	A	0.00	12.37	23.16
7	A	0.30	10.69	31.87
6	A	0.50	4.64	33.92
1	A	0.17	5.47	34.93
8	A	0.00	6.30	27.00
4	A	0.00	0.32	38.59
	A	0.14	6.33	31.67
		0.20	5.56	36.57

TABLE 8G

Colt part II (student talk): Information gap
Mean percentage of observed time for each category

SCH	TYPES	PREDICT	UNPREDIC	DISP REQ	INFO REQ
5	E	8.69	28.30	0.00	3.60
2	E	9.58	5.53	0.51	1.30
	E	9.16	16.18	0.27	2.38
3	A	11.28	2.56	0.70	0.56
7	A	5.83	4.82	0.00	2.45
6	A	14.89	2.06	0.37	2.51
1	A	10.13	11.34	0.00	0.00
8	A	8.71	4.30	0.00	0.26
4	A	14.86	3.70	0.00	1.09
	A	11.05	5.07	0.16	1.01
		10.49	8.37	0.19	1.42

TABLE 8H

Colt part II (student talk): Sustained speech
Mean percentage of observed time for each category

SCH	TYPE	ULTRAMIN	MINIMAL	SUSTAINED
5	E	10.63	8.60	26.04
2	E	8.53	8.40	11.22
	E	9.51	8.49	18.14
3	A	2.49	10.46	2.82
7	A	13.59	11.15	1.98
6	A	13.82	13.84	1.60
1	A	7.55	9.33	9.79
8	A	11.51	8.96	1.09
4	A	8.13	15.96	4.32
	A	9.32	11.50	3.87
		9.38	10.61	8.10

TABLE 8I

Colt part II (student talk): Reaction to message/code
Mean percentage of observed time for each category

SCH	TYPE	CODE	MESSAGE
5	E	0.95	5.85
2	E	1.06	7.85
	E	1.01	6.92
3	A	1.47	0.17
7	A	0.24	2.42
6	A	0.00	1.96
1	A	0.00	3.17
8	A	1.24	2.38
4	A	1.31	0.92
	A	0.72	1.88
		0.81	3.37

TABLE 8J

Colt part II (student talk): Incorporation of preceding utterances
Mean percentage of observed time for each category

SCH	TYPE	CORRECTION	REPETITION	PARAPHRASE	COMMENT	EXPANSION	CLARIF REQ	ELABOR REQ	UNINTEL
5	E	0.52	2.33	0.00	1.65	2.24	0.25	0.00	14.67
2	E	0.43	3.12	0.88	3.02	2.56	1.53	0.00	9.64
	E	0.47	2.75	0.47	2.38	2.41	0.93	0.00	11.99
3	A	0.00	0.56	0.00	0.00	0.00	0.00	0.00	7.37
7	A	0.24	2.30	0.67	2.53	1.05	3.26	0.00	5.13
6	A	0.54	3.11	0.25	1.80	1.56	0.43	0.37	4.65
1	A	0.00	1.88	0.00	0.76	1.27	0.93	0.00	8.25
8	A	0.92	1.38	0.00	3.75	0.49	0.97	0.00	5.43
4	A	0.29	3.83	0.31	1.58	0.61	0.52	0.15	10.17
	A	0.34	2.15	0.17	1.74	0.82	0.93	0.08	6.97
		0.38	2.33	0.26	1.93	1.29	0.93	0.05	8.46

TABLE 8K

Colt part II (student talk): Discourse initiation
Mean percentage of observed time for each category

SCH	TYPE	DISC INIT
5	E	0.00
2	E	1.26
	E	0.67
3	A	0.49
7	A	0.00
6	A	0.50
1	A	0.00
8	A	0.00
4	A	0.61
	A	0.26
		0.38

TABLE 8L

Colt part II (student talk): Restriction of linguistic form
Mean percentage of observed time for each category

SCH	TYPE	RESTRICTED	UNRESTRICTED
5	E	8.13	28.82
2	E	4.09	3.09
	E	5.98	15.12
3	A	14.60	0.28
7	A	7.25	0.46
6	A	12.86	4.64
1	A	9.41	2.45
8	A	8.40	1.14
4	A	16.81	3.64
	A	11.57	2.15
		9.91	6.00

TABLE 9A

Colt part II (teacher talk): Use of target language
Proportion of superordinate category

SCH	TYPE	OFF TASK	L1	L2
5	E	04	10	89
2	E	03	00	99
	E	03	03	96
3	A	02	04	95
7	A	05	00	99
6	A	01	00	99
1	A	04	09	90
8	A	03	08	91
4	A	03	00	100
	A	03	04	95
		03	04	95

TABLE 9B

Colt part II (teacher talk): Information gap
Proportion of superordinate category

SCH	TYPE	PREDICT	UNPREDIC	DISPL REQ	INFO REQ
5	E	11	88	28	71
2	E	04	95	43	56
	E	06	93	35	64
3	A	40	59	13	86
7	A	03	96	00	100
6	A	08	91	42	57
1	A	00	100	19	80
8	A	06	93	09	90
4	A	41	58	53	46
	A	20	79	23	76
		16	83	27	72

TABLE 9C

Colt part II (teacher talk): Sustained Speech
Proportion of superordinate category

SCH	TYPE	MINIMAL	SUSTAINED
5	E	58	41
2	E	51	48
	E	53	46
3	A	75	24
7	A	54	45
6	A	78	22
1	A	63	36
8	A	73	26
4	A	65	34
	A	68	31
		64	35

TABLE 9D

Colt part II (teacher talk): Reaction to message/code
Proportion of superordinate category

SCH	TYPE	CODE	MESSAGE
5	E	10	89
2	E	33	66
	E	23	76
3	A	85	14
7	A	53	46
6	A	68	31
1	A	59	40
8	A	60	40
4	A	75	25
	A	64	35
		53	46

TABLE 9E

Colt part II (teacher talk): Incorporation of preceding utterances
Proportion of superordinate category

SCH	TYPE	CORRECTION	REPETITION	PARAPHRASE	COMMENT	EXPANSION	CLARIF REQ	ELABOR REQ	UNINTEL
5	E	09	25	03	20	17	14	09	03
2	E	07	13	10	16	09	19	23	04
	E	07	17	07	17	12	17	18	04
3	A	18	18	01	39	08	07	07	00
7	A	10	23	10	20	21	07	07	01
6	A	09	12	04	32	09	05	25	00
1	A	08	15	05	29	08	12	20	01
8	A	10	23	02	24	06	15	16	01
4	A	31	12	00	31	08	03	11	04
	A	13	17	04	29	10	09	16	01
		11	17	05	26	10	11	16	02

TABLE 9F

Colt part II (student talk): Use of target language
Proportion of superordinate category

SCH	TYPE	OFF TASK	L1	L2
5	E	01	07	92
2	E	00	06	93
	E	00	06	93
3	A	00	28	71
7	A	01	16	83
6	A	00	05	94
1	A	00	17	82
8	A	00	17	82
4	A	00	00	99
	A	00	13	86
		00	11	88

TABLE 9G

Colt part II (student talk): Information gap
Proportion of superordinate category

SCH	TYPE	PREDICT	UNPREDIC	DISP REQ	INFO REQ
5	E	46	53	00	100
2	E	65	34	16	83
	E	56	43	05	94
3	A	71	28	57	42
7	A	53	46	00	100
6	A	83	16	11	88
1	A	62	37	00	00
8	A	66	33	00	100
4	A	79	20	00	100
	A	70	29	16	83
		66	33	12	87

TABLE 9H

Colt part II (student talk): Sustained speech
Proportion of superordinate category

SCH	TYPE	ULTRAMIN	MINIMAL	SUSTAINED
5	E	43	31	24
2	E	37	33	28
	E	40	32	26
3	A	18	70	11
7	A	48	45	06
6	A	45	51	03
1	A	39	48	11
8	A	54	41	03
4	A	26	58	14
	A	40	51	08
		40	45	14

TABLE 9I

Colt part II (student talk): Reaction to message/code
Proportion of superordinate category

SCH	TYPE	CODE	MESSAGE
5	E	13	86
2	E	13	86
	E	13	86
3	A	87	12
7	A	14	85
6	A	00	100
1	A	00	100
8	A	52	47
4	A	37	62
	A	34	65
		24	76

TABLE 9J

Colt part II (student talk): Incorporation of preceding utterances
Proportion of superordinate category

SCH	TYPE	CORRECTION	REPETITION	PARAPHRASE	COMMENT	EXPANSION	CLARIF REQ	ELABOR REQ	UNINTEL
5	E	07	29	00	25	33	03	00	22
2	E	03	36	06	21	18	13	00	19
	E	04	34	04	22	22	10	00	20
3	A	00	100	00	00	00	00	00	22
7	A	03	23	07	30	11	23	00	14
6	A	09	34	03	15	28	03	06	09
1	A	00	35	00	20	30	15	00	24
8	A	12	24	00	39	09	15	00	14
4	A	02	48	05	27	08	05	02	23
	A	06	34	03	26	16	11	02	18
		05	34	03	25	18	10	01	19

TABLE 9K

Colt part II (student talk): Discourse initiation
Proportion of superordinate category

SCH	TYPE	DISC INIT
5	E	00
2	E	03
	E	02
3	A	01
7	A	00
6	A	01
1	A	00
8	A	00
4	A	01
	A	00
		01

TABLE 9L

Colt part II (student talk): Restriction of linguistic form
Proportion of superordinate category

SCH	TYPE	RESTRICTED	UNRESTRICTED
5	E	43	56
2	E	57	42
	E	47	52
3	A	97	02
7	A	94	05
6	A	69	30
1	A	75	24
8	A	85	14
4	A	78	21
	A	81	18
		73	26

TABLE 10

Colt part II: Time coded, number of turns, proportion of teacher to student turns

School	Time Coded	Total Turns	Teacher Turns	Student Turns	Proportion of turns Teacher	Student
1	44	348	199	149	.57	.43
2	49	426	249	177	.58	.42
3	34	211	121	90	.57	.43
4	38	310	152	158	.49	.51
5	43	237	107	130	.45	.55
6	33	264	152	112	.58	.42
7	27	225	135	90	.60	.40
8	42	373	222	151	.60	.40
Total	310	2394	1337	1057	.56	.44

Table 11

Selected experiential activities in Type A and Type E classes (activities performed 'quite often' or 'very often' during the year, expressed as a percentage of the total number of respondents)

		A (N=6)	E (N=2)
1.	Students read texts in French, concentrating on getting the main ideas without necessarily understanding each individual word.	33	100
2.	Students participated in teacher-led discussions with the whole class on topics of personal and/or general interest.	50	100
3.	Students practised conversational skills by working in pairs or groups to discuss topics of personal or general interest, with the emphasis on the sharing of ideas without worrying too much about making mistakes.	17	50
4.	Students listened to spoken French materials that were not specifically produced for FSL learners, e.g., taped radio shows, news broadcasts, weather forecasts, etc.	0	50
5.	Students were taught aspects of paragraph or text structure, e.g., adverbs such as first, next; clause and sentence or logical connectors (although, however, therefore) or other features relevant to the organization of ideas in a text.	17	100
6.	Students were taught the social and cultural rules of language, e.g., how to make polite requests, address strangers, express opinions, disagree politely, etc.	50	100
7.	In class, the teacher tried to focus on the meaning of what the students said, and not so much on whether their use of language was accurate.	50	50
8.	In correcting written work, the teacher focused on stylistic appropriateness and/or the logical organization of text.	17	100
9.	For homework, the teacher assigned creative writing tasks, e.g., compositions, letters, reports, projects, etc., with no models or guidance provided.	0	50
10.	For homework, the teacher assigned creative oral tasks, e.g., student-prepared oral presentations, role-play situations, oral reports, etc.	50	50

Table 12

Selected analytic activities in Type A and Type E classes (activities performed 'quite often' or 'very often' during the year, expressed as a percentage of the total number of respondents)

		A (N=6)	E (N=2)
1.	Students did single-sentence, fill-in-the-blank exercises.	100	50
2.	Students did guided writing tasks based on pictures, picture series, diagrams, etc.	17	0
3.	Students read French texts, and the teacher asked questions to see if students understood specific points in the text.	100	100
4.	Students practised by repeating words and/or sentences after hearing teacher/tape, or by doing oral substitution or transformation exercises where accuracy is expected.	33	0
5.	Students listened to extended spoken texts, recorded for FSL learners and delivered at a reduced speed, carefully articulated, etc.	17	0
6.	Students were provided with rules and explanations in a way which focused directly on the formal features of language.	83	100
7.	Students did grammar and/or vocabulary exercises either orally or in written form.	100	100
8.	In correcting written work, the teacher focused on spelling, grammar and/or use of vocabulary	100	100
9.	In correcting oral work, the teacher focused on pronunciation, grammar, and/or use of vocabulary	100	100
10.	For homework, the teacher assigned grammar and vocabulary exercises in a single-sentence format	100	100

Table 13

Pre-test and post-test scores and adjusted post-test means for Group E and Group A on main proficiency measures (expressed as percentage of total possible)

(a) = pre-test scores, (b) = post-test scores, (c) = adjusted post-test means.

	ALL STUDENTS						INTERVIEWED STUDENTS					
	GROUP E			GROUP A			GROUP E			GROUP A		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Written tests												
Grammar	52	54	53*	51	56	57*	55	56	55	52	58	59
Discourse	71	65	69	77	72	71	68	66	67	72	76	75
Sociolinguistic	34	32	31	28	33	34	31	28	28	28	30	30
Listening comp.	60	70	69	57	69	70	59	70	69	57	68	68
Oral tests												
Grammar							50	53	54	53	53	53
Discourse							6	6	6	5	6	7
Sociolinguistic							0.1	0.5	0.4	0.1	0.7	0.7
Indices of non-comp.							7	14	15	9	12	12
Sentence rep. (scalar)							40	46	49	46	51	47
Sentence rep. (count)							24	33	38	31	40	34

*Indicates significance at the .06 level

Table 14

Pre-test and post-test scores and adjusted post-test means for Group E and Group A on detailed proficiency measures (expressed as percentage of total possible)

(a) = pre-test scores, (b) = post-test scores, (c) = adjusted post-test means.

	ALL STUDENTS						INTERVIEWED STUDENTS					
	GROUP E			GROUP A			GROUP E			GROUP A		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Grammatical written												
Subcategorization	51	53	52	49	54	54	53	57	56	49	54	55
Gender	56	57	55	50	53	55	67	64	61	54	52	54
Agreement	54	57	58	59	64	63	52	52	54	62	67	65
Grammatical oral												
Subcategorisazion							54	49	49	57	58	58
Gender							71	76	77	76	75	74
Agreement							95	93	93	93	95	95
Sociolinguistic												
Formal letter	65	62	62**	61	66	66**	64	59	59	62	64	64
Informal note	49	47	47**	48	50	50**	48	48	48	48	50	50
Discourse												
Formal letter	59	53	53	64	62	61	56	57	57	62	65	65
Informal note	81	70	71	89	83	82	79	73	73	81	85	85

**Indicates significance at the .01 level

Table 15

Correlations among adjusted post-test means for proficiency measures and COLT orientation score

	COLT	GW	DW	SW	LC	GO	DO	SO
Grammatical written	-.05							
Discourse written	.11	.83**						
Sociolinguistic written	-.10	.83**	.82**					
Listening comprehension	-.05	.91**	.90**	.73*				
Grammatical oral	.27	.56	.34	.45	.35			
Discourse oral	-.08	.47	.47	.31	.56	.62*		
Sociolinguistic oral	-.16	.49	.35	.17	.61	.51	.89**	
Sentence repetition	-.15	.35	.32	.26	.27	.58	.67*	.43

* Indicates significance at the .05 level

** Indicates significance at the .01 level

Table 16

Mean and range for each category of COLT Part I (expressed as percentage of observed time)

COLT COLUMN	MEAN	RANGE
3	50.86	25.95 - 67.48
4	27.01	14.14 - 43.46
5	0.72	0.00 - 2.85
6	0.54	0.00 - 4.34
7	1.75	0.00 - 14.04
8	16.54	2.46 - 37.78
9	2.55	0.00 - 18.57
10-11	8.84	5.68 - 14.06
12	53.22	23.89 - 71.42
13	0.00	0.00 - 0.00
14	0.58	0.00 - 2.70
15	0.12	0.00 - 0.97
16-19	0.14	0.00 - 0.76
20-24	15.10	3.33 - 26.98
25-29	20.07	3.80 - 37.43
30	84.03	44.65 - 100.00
31	10.80	0.00 - 38.16
32	5.16	0.00 - 24.13
33	46.69	35.71 - 65.84
34	18.30	11.83 - 23.73
35	19.52	15.84 - 23.29
36	15.04	0.00 - 26.90
37	0.43	0.00 - 3.40
38	39.49	4.92 - 53.67
39	37.24	5.00 - 68.96
40	0.90	0.00 - 5.00
41	6.78	0.00 - 11.97
42	50.13	33.99 - 68.36
43	1.15	0.00 - 7.89
44	2.31	0.00 - 9.65
45	11.09	0.00 - 36.64

Table 17

Correlations of COLT Part I categories with adjusted post-test means on proficiency measures

COLT COLUMN		GRAM WRIT	DISC WRIT	SOC WRIT	LIST COMP	GRAM ORAL	DISC ORAL	SOC ORAL	SENT REP
3	Teacher/whole class	.37	.39	.57	.36	.52	.40	.20	.15
4	Student/whole class	-.61	-.60	-.27	-.80*	-.31	-.62	-.79*	-.06
5	Choral	-.50	-.39	-.42	-.47	-.75*	-.58	-.54	-.15
6-7	Group								
8	Individual/same	.44	.51	.06	.53	-.01	.18	.35	.14
9	Individual/different								
10-11	Management	.36	.16	.53	.23	.27	-.13	-.07	-.42
12	Form	.31	.42	.10	.57	-.11	.60	.69	.28
13-15	Function/Discourse/Socio.								
16-19	Narrow range								
20-24	Limited range	-.37	-.64	-.30	-.67	.26	-.34	-.41	.13
25-29	Broad range	-.09	-.06	.10	-.32	.30	-.39	-.58	-.12
30	Teacher control	.05	-.12	.38	-.11	.13	.09	-.07	.27
31	Joint control	-.01	.08	-.44	.14	-.15	.05	.28	-.13
32	Student control								
33	Listening	-.17	.05	.23	-.15	.13	-.05	-.31	-.24
34	Speaking	-.68	-.78*	-.55	-.77*	-.52	-.65	-.51	-.51
35	Reading	.35	.21	.25	.22	-.18	-.21	-.17	.34
36	Writing	.44	.29	-.06	.54	.17	.50	.74*	.36
37	Other								
38	Minimal text	.01	-.25	-.09	-.08	-.07	.11	.20	.37
39	Extended text	.16	.34	.16	.19	.42	.16	.02	-.02
40	Audio								
41	Visual	.48	.17	.41	.41	.44	.56	.67	.26
42	L2 material	.28	.18	.38	.16	.50	.57	.32	.87**
43	L1 material								
44	L1-adapted								
45	Student made	.07	.10	-.27	.11	.18	-.04	.10	-.20

* Indicates significance at the .05 level
** Indicates significance at the .01 level

Table 18

Mean and range for each category of COLT Part II (expressed as percentage of observed time)

COLT COLUMN	MEAN	RANGE		MEAN	RANGE
1	3.00	1.70 - 4.61	19	0.21	0.00 - 0.72
2	2.19	0.00 - 5.40	20	5.95	0.33 - 12.37
3	97.81	94.63 - 100.00	21	94.05	87.63 - 99.67
4	1.57	0.00 - 4.36	22	0.36	0.00 - 1.27
5	11.69	4.95 - 18.68			
6	2.25	0.00 - 4.25	23	10.50	5.84 - 14.89
7	6.59	3.01 - 9.62	24	7.83	2.07 - 28.31
			25	0.20	0.00 - 0.70
8	26.77	16.10 - 36.41	26	1.48	0.00 - 3.61
9	25.34	17.41 - 34.06			
			27	9.54	2.50 - 13.82
10	3.17	0.51 - 6.59	28	10.84	8.41 - 15.96
11	2.94	0.59 - 4.83	29	7.36	1.10 - 26.04
12	3.70	1.99 - 7.42	30	10.20	4.09 - 16.81
13	4.82	3.08 - 6.68	31	5.57	0.28 - 28.83
14	1.91	0.00 - 5.07			
15	8.39	4.64 - 13.19	32	0.79	0.00 - 1.47
16	4.51	1.70 - 9.93	33	3.09	0.17 - 7.86
17	3.11	0.85 - 6.37			
18	4.56	1.49 - 9.19	34	0.37	0.00 - 0.93
			35	2.32	0.56 - 3.83
			36	0.27	0.00 - 0.89
			37	1.89	0.00 - 3.76
			38	1.23	0.00 - 2.56
			39	0.99	0.00 - 3.27
			40	0.07	0.00 - 0.38

Table 19a

Correlations of COLT Part II categories with adjusted post-test means on proficiency measures (teacher interaction)

	COLT COLUMN	GRAM WRIT	DISC WRIT	SOC WRIT	LIST COMP	GRAM ORAL	DISC ORAL	SOC ORAL	SENT REP
1	Off-task	.05	.25	.21	.31	.05	.64	.56	.02
2	L1	.01	-.36	-.08	-.13	.25	-.12	.03	-.21
3	L2	-.01	.36	.08	.13	-.25	.12	-.03	.21
4	Predictable	-.64	-.81*	-.82*	-.64	-.50	-.45	-.16	-.42
5	Unpredictable	-.07	-.23	-.30	-.19	.28	-.01	.20	-.21
6	Display request	-.54	-.35	-.71	-.35	-.74*	-.45	-.34	-.26
7	Information request	-.22	-.27	-.17	-.10	.29	.53	.51	.01
8	Minimal speech	-.07	-.37	-.18	-.34	.26	-.04	.01	.39
9	Sustained speech	.34	.61	.34	.47	.48	.82*	.58	.63
10	Code reaction	.28	.14	.30	.22	.50	.68	.50	.85**
11	Message reaction	.34	.50	.50	.31	.41	.18	-.10	-.24
12	Correction	-.65	-.62	-.70	-.54	-.70	-.41	-.20	-.46
13	Repetition	.25	.07	.50	.08	.47	.05	.01	-.24
14	Paraphrase	.36	.68	.46	.43	.47	.53	.29	.32
15	Comment	-.26	-.51	-.35	-.47	.14	-.04	-.05	.50
16	Expansion	-.13	.11	.27	-.22	.34	.17	-.18	.21
17	Clarif. request	.82*	.83**	.62	.76*	.66	.48	.40	.47
18	Elabor. request	.41	.55	.36	.33	.29	.24	-.02	.75*

* Indicates significance at the .05 level
** Indicates significance at the .01 level

Table 19b

Correlations of COLT Part II categories with adjusted post-test means on proficiency measures (student interaction)

COLT COLUMN		Gram WRIT	DISC WRIT	SOC WRIT	LIST COMP	GRAM ORAL	DISC ORAL	SOC ORAL	SENT REP
19	Off-task	-.46	-.24	-.03	-.52	-.10	-.40	-.71*	-.15
20	L1	-.17	-.42	-.09	-.34	.57	.27	.25	.06
21	L2	.17	.42	.09	.34	-.57	-.27	-.25	-.06
22	Disc. initiation	.04	.20	-.29	.12	-.11	-.02	.07	.12
23	Predictable	-.43	-.41	-.53	-.43	-.61	-.46	-.40	.09
24	Unpredictable								
25	Display request	-.16	-.28	-.51	-.29	.34	.10	.16	.33
26	Information request	-.47	-.16	-.07	-.51	-.21	-.54	-.77*	-.42
27	Ultramin. speech	.25	.43	.70	.12	.03	-.26	-.55	-.02
28	Minimal speech	-.49	-.34	-.31	-.44	-.73*	-.46	-.46	-.19
29	Sustained speech	-.25	-.04	-.21	-.10	-.17	-.18	-.24	-.33
30	Restricted form	-.64	-.76*	-.66	-.63	-.63	-.43	-.28	-.20
31	Unrestricted form								
32	Code reaction	.00	-.27	-.39	.00	-.21	-.35	.10	-.62
33	Message reaction	.35	.60	.28	.44	.22	.13	.02	.03
34	Correction	.50	.32	.55	.26	.07	-.52	-.46	-.24
35	Repetition	-.12	.33	.06	.06	-.65	-.36	-.49	-.18
36	Paraphrase	.26	.62	.28	.37	.08	.23	.11	.05
37	Comment	.76*	.75*	.81*	.64	.22	-.10	-.09	-.12
38	Expansion	.10	.48	.20	.19	.05	-.02	-.27	.09
39	Clarif. request	.42	.65	.66	.49	.34	.48	.30	.08
40	Elabor. request								

* Indicates significance at the .05 level

Table 20

Correlations of COLT combined categories with adjusted post-test means on proficiency measures

	GRAM WRIT	DISC WRIT	SOC WRIT	LIST COMP	GRAM ORAL	DISC ORAL	SOC ORAL	SENT REP
1(a) whole class activity (b) group/individual activity (I:3-5 vs. I:6-9)	.01 -.01	.03 -.03	.41 -.41	-.11 .11	.31 -.31	.02 -.02	-.28 .28	.11 -.11
2(a) focus on form (b) limited/broad range of reference (I:12 vs. I:20-29)	.31 -.25	.42 -.37	.10 -.10	.57 -.54	-.11 .32	.60 -.41	.69 -.56	.28 .00
3(a) teacher control (b) student/joint control (I:30 vs. I:31-32)	.05 -.05	-.12 .12	.38 -.38	-.11 .11	.13 -.13	.09 -.09	-.07 .07	.27 -.27
4(a) speaking/sustained speech (b) writing/extended text (I:34 + II:9 vs. I:36 + I:39)	-.15 .35	.11 .46	-.05 .13	-.07 .41	.16 .49	.46 .38	.28 .33	.34 .13
5(a) L2 materials (b) L1/adapted/student-made materials (I:42 vs. I:43-45)	.28 .19	.18 .17	.38 -.10	.16 .12	.50 .34	.57 -.10	.32 -.01	.87 -.15
6(a) predictable/display request (b) unpredictable/inform request (II:4 + 6 + 23 + 25 vs. II:5 + 7 + 24 + 26)	-.54 .55	-.46 .62	-.45 .27	-.60 .51	.28 .29	.21 .11	.03 .08	.29 .34
7(a) reaction to code (b) reaction to message (II:10 + 32 vs. II: 11 + 33)	-.54 .45	-.43 .67	-.11 .34	-.57 .72*	-.56 .09	-.75* .55	-.84* .50	-.58 .20
8(a) restricted incorporation (b) expanded incorporation (II:12-14 + 34-36 vs. II:15-18 + 37-40)	.67 .28	.22 .37	.44 .65	.52 .37	.60 .08	.50 .33	.67 .17	.37 -.03

* Indicates significance at the .05 level

Table 21

Average rank of schools based on adjusted post-test means for each measure

SCHOOL	WRITTEN MEASURES	ORAL MEASURES	TOTAL
2	7.1	6.3	6.7
8	7.5	5.3	6.4
1	5.6	7.0	6.3
7	5.8	5.0	5.4
6	3.5	3.5	3.5
3	1.3	5.5	3.4
4	3.0	2.0	2.5
5	2.3	1.5	1.9

APPENDIX A

The Ontario Institute for Studies in Education
252 Bloor Street West, Toronto, Ontario M5S 1V6 Tel. 923-6641

Modern Language Centre

FRENCH CONTACT QUESTIONNAIRE Form 1

In this questionnaire, you are asked a few questions about your contact with French outside the classroom.

Take your time responding to the questions and try to answer them as honestly as possible. There are no right or wrong, or good and bad answers. Also remember, this study is confidential and the results do not affect your school marks in any way.

Affiliated with the University of Toronto

1. **Name:** ____

2. **School:** ____

3. **Date of Birth: Month:**____________________ **Year:** __________

4. **Place of Birth: Country:**____________________**City (Town):**__________

   (4a.) If you were born outside Canada, when did you come to Canada? Year:______

5. **Have you ever lived in a community where there were many francophone speakers? Yes:**__________ **No:** __________

   (5a.) If yes, where: _______________ **(5b.) For how long:** ____________

6. **For how many years (or months) have you lived in Toronto?** _______________

7. **Where did you start learning French?**

   Country: ____________________

   Place: ____________________

8. **In what grade did you start learning French in school? (If you don't remember the exact grade, write down the approximate grade level.)**

   Grade: __________

9. **Do you know any language(s) other than English and French?**

   Yes: __________ **No:** __________

   If yes:

   (9a.) Which language(s)? 1. ______________________________

   2. ______________________________

   3. ______________________________

   4. ______________________________

(9b.) How well do you know each of these languages? (Please put check mark into the appropriate box)

	A little	Fairly Well	Very Well
1. ______________________			
2. ______________________			
3. ______________________			
4. ______________________			

10. How often is each of the following languages spoken in your home?

	Never	Hardly ever	Sometimes	Often	Most or All of the Time
English	( )	( )	( )	( )	( )
French	( )	( )	( )	( )	( )
Other (please specify)					
______________________	( )	( )	( )	( )	( )
______________________	( )	( )	( )	( )	( )

11. In general, how often have you had the opportunity to use French in some way outside of school during the past five years?

Never	Hardly Ever	Sometimes	Often	Extremely Often
( )	( )	( )	( )	( )

12. How much time have you spent altogether in a French-speaking country (or province) in the past five years?

None	Up to 2 Weeks	2½ to 4 Weeks	4½ to 6 Weeks	More than 6 weeks
( )	( )	( )	( )	( )

13. In the past five years, have you participated in any French activities outside of school, e.g. summer camps?

Yes: ______ No: _____

If Yes: (13a.) Please specify activity: ______________________________

(13b.) How much time was involved in total?

Up to 2 Weeks	2½ - 4 Weeks	4½ - 6 Weeks	More than 6 Weeks
( )	( )	( )	( )

THE NEXT FEW QUESTIONS WILL BE ABOUT YOUR PRESENT CONTACT WITH FRENCH OUTSIDE OF YOUR REGULAR FRENCH CLASSES AT SCHOOL.

14. Since the beginning of this school year, how often have you done each of the following in addition to your school work?

	Never	Hardly Ever	Sometimes	Frequently	Daily
a. Talked in French with your friends outside of school					
b. Talked with native speakers of French					
c. Listened to French music					
d. Listened to spoken French on the radio					
e. Watched French T.V.					
f. Been to see French movies					
g. Read French newspapers or magazines					
h. Read brochures and pamphlets in French					
i. Read French advertisements or labels on packages					
j. Read French books					
k. Written letters or notes in French					
l. Written other texts in French, e.g. stories					

15. Are you presently taking any other French courses in addition to the French classes at school?

Yes: ____________ No: ____________

16. Are you presently involved in any extracurricular activites in French, e.g. a French club?

Yes: ____________ No: ____________

17. How much time do you usually spend on your French homework per week?

Less than ½ hour per per week	½ to 1 hour per week	1½ to 2 hours per week	2½ - 3 hours per week	3½ - 4 hours per week	more than 4 hours per week
( )	( )	( )	( )	( )	( )

FRENCH CONTACT QUESTIONNAIRE - FORM 2

In this questionnaire, you are asked a few questions about your contact with French and other languages you may know.

Take your time responding to the questions and try to answer them as honestly as possible. There are no right or wrong, or good and bad answers. Also remember, this study is confidential and the results do not affect your school marks in any way.

1. Name: ____

2. School: ____

3. Date of Birth Month: ______________ Year: __________

4. Have you ever been enrolled in a French Immersion Program? Yes ___ No ___

   If YES: In what grade did you start the program? ______________

   In what grade did you leave the program? ______________

5. Have you ever been enrolled in an Extended French Program? Yes ___ No ___

   If YES: In what grade did you start the program? ______________

   In what grade did you leave the program? ______________

6. Have you ever studied any language(s) other than English or French at school? Yes ___ No ___

   If YES:

   (a) Which language(s)?

   (1) ____________________

   (2) ____________________

   (3) ____________________

   (4) ____________________

   (b) How well do you know each of these languages? (Please circle the appropriate number)

	a little	fairly well	very well
(1) ____________	1	2	3
(2) ____________	1	2	3
(3) ____________	1	2	3
(4) ____________	1	2	3

7. Are any languages other than English or French spoken regularly in your home? Yes ____ No ____

If YES:

(a) Please list the languages each family member uses.

(1) Yourself ______________________________

(2) Father ______________________________

(3) Mother ______________________________

(4) Brother ______________________________

(5) Sister ______________________________

(6) Others (please specify) ______________________________

____

(b) How well do you understand each of these languages?
(List the language(s) and circle the appropriate number.)

		a little	fairly well	very well
(1)	____________	1	2	3
(2)	____________	1	2	3
(3)	____________	1	2	3
(4)	____________	1	2	3

(c) How well do you speak each of these languages?
(List the language(s) and circle the appropriate number.)

		a little	fairly well	very well
(1)	____________	1	2	3
(2)	____________	1	2	3
(3)	____________	1	2	3
(4)	____________	1	2	3

9. Please indicate how well you feel you know English by circling the appropriate number for each line on the chart below.

	not at all	with some difficulty	fairly well	quite well	extremely well
I understand English	1	2	3	4	5
I speak English	1	2	3	4	5
I read English	1	2	3	4	5
I write English	1	2	3	4	5

10. Please indicate how well you feel you know French by circling the appropriate number for each line on the chart below.

	not at all	with some difficulty	fairly well	quite well	extremely well
I understand English	1	2	3	4	5
I speak English	1	2	3	4	5
I read English	1	2	3	4	5
I write English	1	2	3	4	5

THE NEXT FEW QUESTIONS WILL BE ABOUT YOUR PRESENT CONTACT WITH FRENCH OUTSIDE OF YOUR REGULAR FRENCH CLASSES AT SCHOOL.

11. Since the beginning of this school year, how often have you done each of the following in addition to your school work and assigned homework?

		never	hardly ever	some-times	frequently	daily
a.	spoken French with your friends outside of school	1	2	3	4	5
b.	spoken French with native speakers of French	1	2	3	4	5
c.	listened to French music	1	2	3	4	5
d.	listened to spoken French on the radio	1	2	3	4	5
e.	watched French on T.V.	1	2	3	4	5

11. (continued)

		never	hardly ever	some-times	frequently	daily
f.	been to see French movies	1	2	3	4	5
g.	read French newspapers or magazines	1	2	3	4	5
h.	read brochures or pamphlets in French	1	2	3	4	5
i.	read French advertisements or labels on packages	1	2	3	4	5
j.	read French books	1	2	3	4	5
k.	written letters or notes in French	1	2	3	4	5
l.	written other texts in French, e.g. stories	1	2	3	4	5

12. Are you presently taking any other French courses in addition to the French classes at school? Yes ____ No ____

If YES:

Please indicate what course and how often you have classes.

____

13. Are you presently involved in any extracurricular activities in French, e.g. a French club? Yes ____ No ____

14. How much time do you usually spend on your French homework, per week?

____

____

APPENDIX B

The Ontario Institute for Studies in Education
252 Bloor Street West, Toronto, Ontario M5S 1V6 Tel. 923-6641

Modern Language Centre

GME ____________________

____

____

School: ____

Name: ____

GRAMMAR TEST

This is a multiple choice grammar test. For each blank there are three choices given. Only one is correct.

Choose the correct answer and circle the letter beside it.

Example:

Les livres ______________ Jean sont sur le bureau.

a) au
b) de
c) par

Affiliated with the University of Toronto

1. Il y a beaucoup __________________ bicyclettes dans la rue.

   a) de

   b) des

   c) les

2. Est-ce que tu vas __________________ spectacle ce soir?

   a) à la

   b) au

   c) chez le

3. Pendant que Marie lit, Jacques et Paul ________________ la télévision.

   a) regardaient

   b) ont regardé

   c) regardent

4. Antoine _________________ ses lunettes en faisant du ski.

   a) a cassé

   b) est cassé

   c) s'est cassé

5. Marie s'est dirigée __________________ carré St. Louis.

   a) pour le

   b) au

   c) vers le

6. Ils sont au Canada ____________________ déjà trois mois.

a) durant

b) pour

c) depuis

7. Marie était bien ____________________ aujourd'hui.

a) habiller

b) habillée

c) habillé

8. Mme Beauséjour regardait par la fenêtre quand ils ____________________ dans la chambre.

a) entreront

b) entrent

c) entrèrent

9. Elle a beaucoup ____________________ les blouses vertes.

a) aimé

b) aimée

c) aimées

10. Jean et Marie partent ____________________ vacances la semaine prochaine.

a) aux

b) pour

c) en

11. Jean a mangé le gâteau que tu m' ____________________.

a) apporterais

b) apportes

c) as apporté

12. Je m'attends ____________________ ce qu'elle arrive bientôt.

a) à

b) pour

c) que

13. Ils pensent ____________________ toi tous les jours.

a) de

b) à

c) en

14. Je leur ai ____________________ des verres en cadeau.

a) données

b) donnés

c) donné

15. Les enfants de l'école Ste-Marie ____________________ l'autobus tous les jours.

a) prend

b) prends

c) prennent

16. À ta place, je le ferais tout de suite au cas où quelqu'un ________________.

a) arriverait

b) arrivait

c) arrivera

17. Laquelle des deux rédactions est la mieux ____________________?

a) écrit

b) écrites

c) écrite

18. Je _____________________ donnerai un livre pour sa fête.

a) le

b) la

c) lui

19. Ils ont fini ______________________ travaux à dix heures.

a) ses

b) leurs

c) leur

20. Voulez-vous ____________________ la table de vos livres?

a) débarrassez

b) débarrassé

c) débarrasser

21. Est-ce qu'ils racontent souvent ____________________ histoires?

a) telles

b) de telles

c) de tels

22. ______________________ de ces livres veux-tu garder?

a) Lesquels

b) Desquels

c) Lesquelles

23. Ils ont acheté ____________________ meuble chez l'antiquaire.

a) un vieil

b) une vieille

c) un vieux

24. La semaine passée, j'ai parlé à Marie comme si elle __________________ une adulte.

a) était

b) serait

c) a été

25. Les États-Unis se trouvent entre le Canada et ____________ Mexique.

a) la

b) le

c) --

26. Nous sommes partis de Toronto pour aller ______________ Montréal.

a) en

b) au

c) à

27. Si j'avais le temps, je ____________________ avec toi.

a) jouerais

b) jouerai

c) jouais

28. Lorsque je suis fatigué je rentre ____________________.

a) chez moi

b) à ma maison

c) à moi

29. Elles sont parties sans ____________________ à personne.

a) avoir parler

b) avoir parlé

c) avoir parlées

30. Les livres sont sur ____________________ bureau.

a) cet

b) cette

c) ce

31. J'ai étudié à Paris ________________ trois années.

a) pour

b) pendant

c) --

32. Depuis qu'il est parti nous ________________ seuls.

a) resterions

b) sommes restés

c) étions restés

33. Dépêchez-vous sinon vous ________________ en retard.

a) seriez

b) serez

c) êtes

34. Nous sommes assis ________________ vous.

a) devant

b) envers

c) arrière

35. J'évite ________________ prendre l'autobus pendant les heures de pointe.

a) de

b) pour

c) à

36. Si tu n'étudies pas tu ne ____________________ pas tes examens.

a) réussirais

b) réussiras

c) réussis

37. La dame à qui j'ai ____________________ n'est pas chez elle.

a) téléphoné

b) téléphonée

c) téléphoner

38. Il faut ____________________ ce genre d'erreur.

a) évitez

b) éviter

c) évité

The Ontario Institute for Studies in Education
252 Bloor Street West, Toronto, Ontario M5S 1V6 Tel. 923-6641

Modern Language Centre

Name:

School:

Writing Task 1

Note

You forgot that your parents are expecting visitors tonight. Your room is in a mess and your things are all over the living-room. Your mother is not very pleased with you. When you come home from school, she is not there, but she has left you a note on the kitchen table. It is not difficult to imagine what she is asking you to do.

Write the note in the way you think your mother would have written it.

Affiliated with the University of Toronto

Note

Tu as oublié que tes parents ont de la visite ce soir. Ta chambre est très en désordre et tes affaires traînent partout dans le salon. Ta mère n'est pas contente de toi. Quand tu rentres de l'école, elle n'est pas là, mais elle t'a laissé une note sur la table de cuisine. Ce n'est pas difficile d'imaginer ce qu'elle te demande de faire. Ecris la note comme tu penses que ta mère l'aurait écrite.

The Ontario Institute for Studies in Education
252 Bloor Street West, Toronto, Ontario M5S 1V6 Tel. 923-6641

Modern Language Centre

Name:

School:

Writing Task 2

A letter

Imagine that your family has rented a house in the country for the month of August. In the garage, which is locked, you see a beautiful 10-speed bicycle.

Write a letter to the landlord and try to convince him to let you use the bicycle.

Use the space provided on the next page.

Imagine-toi que ta famille loue une maison à la campagne pendant le mois d'août. Dans le garage, qui est barré, tu vois une belle bicyclette dix-vitesses.

Ecris une lettre au propriétaire de la maison pour le convaincre de te donner la permission d'utiliser la bicyclette.

Sers-toi de l'espace ci-dessous:

Cher monsieur, ______________________________

FRENCH LISTENING COMPREHENSION TEST

STUDENT BOOKLET

Name: ANSWERS -

School: ____________________

Grade: __________

French Listening Comprehension Test

This is a test of your ability to understand spoken French. You will hear a number of conversations, announcements and short broadcasts in French. At the end of each recorded passage you will be asked one or more questions in English about what has been said. For each question you are to select the best answer from among the four choices printed in your test booklet.

Your score will be based on the number of questions you have answered correctly. It is to your advantage to answer every question even though you may not be sure whether your answer is correct. If you make a mistake or wish to change an answer, erase your first answer and then mark your new choice.

Before you listen to each passage you will have an opportunity to read the question and the choices printed in your booklet.

Number 0. Where did this conversation take place?

(A) In the kitchen.

[✓ (B)] At the market.

(C) In a restaurant.

(D) On a farm.

Number 1. What are the people advised to do?

(A) They must rest.

[✓ (B)] They should not go out.

(C) They must abandon their homes.

(D) They must hurry up.

Number 2. What is being talked about?

(A) That our team has lost the game.

(B) That our team is not going to play again this year.

[✓ (C)] That our team is going to play in another city.

(D) That our team is going to stay here.

Number 3. What has happened?

[✓ (A)] The date of the performance has been changed.

(B) An actor has died.

(C) The director has resigned.

(D) A performance has been interrupted.

Number 4. What has happened?

(A) Some explorers have been making drawings.

(B) The photographers have left.

(C) Some animals have been chased into caves.

[✓ (D)] An important discovery has been made.

Number 5. What was this announcement about?

[✓ (A)] A violent storm.

(B) The modernization of a village.

(C) A plane crash.

(D) A battle.

Number 6. What is the main point of this announcement?

[✓ (A)] Carole's sister is going to recite a poem.

(B) Carole is going to recite a poem.

(C) Carole's sister recited a poem last week.

(D) Carole's sister has the flu.

Number 7. What has been announced by CN Rail?

(A) There will be no more train service between Toronto, Montreal and Ottawa.

[✓ (B)] There will be no more restaurants on some trains.

(C) There will be no more restaurants on any trains.

(D) People are advised to go by plane to Montreal and Ottawa instead of taking the train.

Number 8. What was this scene about?

(A) A costume ball.

[✓ (B)] An attempted bank robbery.

(C) A fight.

(D) An accident.

Number 9. Where are the two men?

(A) At the police station.

(B) In a store.

[✓ (C)] Close to the bank.

(D) With the bank manager.

Number 10. What happened to the two men in the end?

(A) Their TV set was stolen.

(B) They got away with millions of francs.

(C) They lost their guns.

[✓ (D)] They were arrested.

Number 11. Who is being interviewed?

(A) A young woman who was born in Italy and raised in Sherbrooke.

(B) A young Italian woman who recently immigrated to Toronto.

[✓ (C) A young woman who was born and raised in Quebec.]

(D) A young woman of Italian background whose parents still live in Italy.

Number 12. Which language or languages do the majority of people in Sherbrooke speak?

[✓ (A) French and English]

(B) French, English and Italian

(C) French only

(D) Italian and French

Number 13. Which of the following statements best summarizes what you have just heard about the Italian and Francophone population in Sherbrooke?

(A) There is almost no difference between the Italian and Francophone way of life.

(B) The only difference is the type of food they eat.

(C) The only difference lies in the language they speak.

[✓ (D)] There are many differences in family life and values.

Number 14. How is the relationship between Sylvana's family and the Francophone population best described?

[✓ (A)] They are getting along very well.

(B) They would prefer to live in Toronto.

(C) They don't want to mix with Francophones.

(D) The parents don't want their children to marry Francophones.

Number 15. What is this conversation about?

(A) The man wants to sell certain items.

(B) The woman wants to know the price of the items which are for sale.

(C) The man wants to buy certain items.

[✓ (D)] The woman wants to sell certain items.

Number 16. What can you do with one of the articles for sale?

(A) Clean the floor.

[✓ (B)] Type a letter.

(C) Heat up food.

(D) Do calculations.

Number 17. Which of the following is true?

(A) They have to be sold together.

(B) One of them costs twice as much as the other.

(C) One of the items is in good condition, the other one needs a lot of repair.

[✓ (D)] The person wants to sell the two items for the same price.

Number 18. Which season is being talked about?

[✓ (A)] Spring

(B) Summer

(C) Fall

(E) Winter

Number 19. What is the weather forecast for that day?

(A) Generally sunny, but slightly cool.

[✓ (B)] Very windy, cool and overcast.

(C) Generally sunny and warm.

(D) Generally sunny and warm with some cloudy periods.

This is the end of the test.

APPENDIX C

ORAL INTERVIEW

(questions in parentheses were asked in the
autumn pretest, others in the spring post-test)

1. Bonjour. Comment ça va? Je m'appelle Susanne. Comment t'appelles-tu?
2. Après l'école, qu'est-ce que tu aimes faire?.
3. Qu'est-ce que tu vas faire aujourd'hui après l'école?
4. Bon, maintenant je veux que tu me racontes en détail ce que tu fais d'habitude le matin avant de venir à l'école.
5. Très bien. Pour changer de sujet, qu'est-ce que tu faisais normalement en avril (septembre), pendant les weekends?
6. As-tu déjà fait un voyage en dehors de Toronto? (As-tu vu les Olympiques à la télévision?)
7. Où es-tu allé? (Qu'est-ce que tu as vu?)
8. Avec qui es-tu allé? (As-tu vu la course entre Mary Decker et Zola Budd?)
9. Qu'est-ce que tu as fait pendant ce voyage? (Raconte-moi ce qui s'est passé?)
10. As-tu des frères et/ou des soeurs? (Si "oui" passe à (11).
11. Qu'est-ce qu'elle(s)/il(s) faisait(aient) ce matin quand tu es parti pour l'école?
12. (à partir du diagramme de la maison). Prends ce plan d'une maison. Imagine que tu es au téléphone et que tu dois décrire la maison à quelqu'un qui ne la voit pas. Par exemple, où se trouve le lit? Y a-t-il une commode dans la chambre? Où est-elle? Où se trouve la télévision? Où exactement dans le salon? Et la lampe? Et le tableau? Dans la salle à manger, il y a un bol de fruits, n'est-ce pas? Où est-il? Et les chaises? Où se trouve le garage? Et la voiture? Et la lumière?
13. Bon. Dis-moi ce que tu ferais si tu avais tout d'un coup beaucoup d'argent? (Si tu gagnais à la loterie?)

14. (à omettre si l'étudiant a répondu à (6) à (9)).

Dis-moi, (a) as-tu déjà eu un accident?

(Si "oui", passe à (15), si "non" passe à (14b).)

(b) As-tu déjà eu très peur?

(Si "oui" passe à (15), si "non" passe à (14c).)

(c) As-tu déjà fait un mauvais rêve?

15. Est-ce que tu peux m'en parler?

16. (à partir du plan du village) Voici un village au Québec. J'ai tracé en rouge le trajet du facteur. Commence par l'école et décris-moi le chemin du facteur en me disant exactement ce qu'il fait et en me nommant tous les endroits où il s'arrête pour laisser du courrier.

17. Maintenant, imagine qu tu es à la bibliothèque où tu étudies pour un examen important. A la table d'à côté se trouve un étudiant de ton âge qui est en train de manger des chips. Il fait beaucoup de bruit, tu sais, "cric, crac". Ça t'empêche de travailler. Tu veux lui dire d'arrêter. Qu'est-ce que tu lui dis? Maintenant, imagine que la personne qui fait le bruit est le directeur de l'école. Qu'est-ce que tu lui dis?

18. Tu es dehors devant l'école. Les voitures passent très vite dans la rue. Tu vois un de tes amis (une de tes amies) sur le trottoir. Il(elle) veut traverser sans regarder autour de lui (elle). Qu'est-ce que tu lui dis pour l'avertir de faire attention aux voitures? Imagine la même scène sauf que cette fois-ci tu vois une vieille dame qui ne voit pas les voitures. Qu'est-ce que tu lui dis pour l'avertir?

19. Maintenant, je veux que tu répètes quatre phrases que je vais te faire écouter. D'abord, tu vas entendre les quatre phrases ensemble. Ensuite, je vais jouer chaque phrase individuellement et tu la répéteras tout de suite après. (On passe à l'enrégistreuse)

APPENDIX D

COLT PART C

INSTRUCTIONAL INPUT

A. Oral production

(a) Students do pronunciation exercises, e.g. practise stress and intonation patterns, and/or learn how to produce and distinguish difficult sounds.

(b) Students practise by repeating words and/or phrases after teacher.

(c) Students begin with model (sentence/minimal text) based on textbook or handout. Then, using the same sentence patterns, they substitute their own vocabulary and ideas.

(d) Students do guided 'conversation' exercises based on model dialogues, simulation, games, role-playing, or any form of controlled practice designed to help students develop appropriate patterns of L2 conversational behaviour.

B. Reading and Listening

(a) Students read or listen to structurally/lexically graded or simplified materials, and designed to extend or reinforce the student's knowledge of grammar and lexis.

(b) Students read or listen to discoursally graded or simplified materials, and designed to extend or reinforce the student's knowledge of discourse organisation.

(c) Students do oral or written comprehension exercises designed to see if they have understood specific points in the text (i.e., comprehension work which focuses on the literal meaning of separate sentences and reinforces knowledge of grammar and lexis).

(d) Students do oral or written comprehension exercises which require them to identify the logical an rhetorical relations in a text (i.e., comprehension work which focuses on patterns of organisation in the text as a whole).

C. Writing

(a) Students do single-sentence substitution or transformation type exercises designed to practise specific aspects of sentence grammar or vocabulary.

(b) Students do guided writing tasks based on pictures, diagrams, fill-in-the-blank model paragraphs, designed to practice vocabulary and specific aspects of discourse.

D. Teaching about the language

1. (a) Teacher gives rules or explanations for grammar, vocabulary and/or phonology, in a way which focuses attention on the formal features of langauge, and involves the use of metalanguage (e.g., 'masculine noun', 'definite article', 'passé composé').

   (b) Teacher explains aspects of paragraph and text structure, e.g. adverbs such as first, next; sentence conjunction (although, therefore, however) and other features relevant to the organisation of ideas in speech or writing.

   (c) Teacher explains the social and cultural rules of language, e.g. how to make polite requests, address strangers, express opinions, disagree politely, etc.

2. (a) Provision of rules or explanations is the main focus of attention, and constitutes a regular classroom activity.

   (b) Rules or explanations are provided incidentally, in order to facilitate the students' performance on some other task.

E. Syllabus design

1. (a) Teaching during the year is based on a planned progression of materials involving the use of structural and/or lexical grading.

   (b) Teaching during the year is based on a planned progression of materials involving the systematic introduction of discourse features or socio-cultural aspects of language use.

2. (a) Teaching during the year is based on a textbook which incorporates simple rules and explanations, and in which regular typographic and diagrammatic conventions are used to focus attention on the point being learned.

3. (a) During the year the teacher encourages the students to use a bilingual dictionary, to compile word-lists, to look for cognates and false cognates, and generally to develop a systematic approach to the learning of vocabulary.

APPENDIX E

TEACHER QUESTIONNAIRE

This questionnaire seeks information about the learning activities of students in the grade 11 core French class which has been observed. It should be emphasized that the researchers who have formulated the questions have no prior commitment to any particular language teaching method. Consequently there is no value judgement intended by any of them. The purpose of this questionnaire is to supplement the information collected during the visits to your class.

Date:

School:

Grade:

Name:

GENERAL INFORMATION

1. How many students are enrolled in the class under observation?__________

2. Approximately how many minutes of French instruction do the students receive per week?__________

3. Which textbook(s) and/or reader(s) have you been using with these students since September 1984? (List them in the space provided.)

4. Since the beginning of this school year have you used any supplementary materials in the class which has been observed?

   Yes:_____ No:_____

   If Yes, please indicate on page 3 the type of supplementary materials and the amount of time per week that you have used each type since September 1984.

Type of supplementary material	less than one hour per week	between 1-2 hours per week	2 hours or more per week

WRITING (during class hours)

5. Since the beginning of the school year, how often have you asked your students to do the following types of writing exercise in class?

	never	rarely	some-times	quite often	very often
(a) copying French language material from blackboard, textbook or other source					
(b) doing fill-in-the-blank type exercises (i) single sentence exercises					
(ii) paragraph exercises					
(c) writing down answers that require two or more sentences					
(d) writing down (i) dictated word lists					
(ii) dictated sentences					
(e) guided writing tasks, based on pictures, picture series, diagrams, etc.					
(f) writing down short (one-sentence) answers to questions					
(g) free composition, e.g. letters, notes, compositions, paragraphs on (i) specified topics					
(ii) self-selected topics					

	never	rarely	some-times	quite often	very often
(h) other (please specify below)					

READING

6. How often did your students read texts in French (either during class time or as homework) that were not specifically written for French as a second language (FSL) learners?

never	rarely	sometimes	quite often	very often
( )	( )	( )	( )	( )

7. When you read texts in French with your students in class, how often did you do each of the following (make a general assessment):

	never	rarely	some-times	quite often	very often
(a) ask questions to see if students understood specific points in the text					
(b) have students concentrate on getting the main ideas without necessarily understanding each individual word					
(c) ask questions which require comprehension of the text as a whole (i.e. not just retrieval of a specific point of information), or ask students to give a summary, identify the main points, etc.					

7. (continued)

	never	rarely	some-times	quite often	very often
(d) explain unknown words or structures					
(e) have students guess the meaning of unknown words					

LISTENING AND SPEAKING

8. Since September 1984, how often have each of the following activities occurred in class (make a general assessment):

	never	rarely	some-times	quite often	very often
(a) students practised by repeating words and/or sentences after hearing teacher/ tape, or by doing oral substitution or transformation exercises where accuracy was expected					
(b) students gave oral presentations based on memorized material from textbooks or other sources					
(c) students provided a summary or identified the main ideas, etc., after listening to a spoken text					
(d) teacher led discussion with the whole class on topics of personal and/or general interest					

8. (continued)

	never	rarely	some-times	quite often	very often
(e) students did pronunciation exercises, e.g. pronouncing and distinguishing difficult sounds					
(f) students provided answers to specific, factual questions based on listening materials					
(g) students listened to extended spoken texts delivered at a normal speed and with natural articulation but still designed for FSL learners					
(h) students gave oral presentations of texts which they created themselves					
(i) students began with "pattern practice" based on textbook. When using the same sentence patterns, they substituted their own vocabulary and ideas					
(j) students listened to extended spoken texts, recorded for FSL learners and delivered at a reduced speed, carefully articulated, etc.					
(k) students practised conversation skills by working in pairs or groups to discuss topics of personal or general interest, with the emphasis on the sharing of ideas without worrying too much about making mistakes					

8. (continued)

	never	rarely	sometimes	quite often	very often
(l) students listened to spoken French materials that were not specifically produced for FSL learners, e.g. taped radio shows, news broadcasts, weather forecasts, etc.					

FOCUS ON LANGUAGE

9. Since the beginning of this school year, how often have you done each of the following in class?

	never	rarely	sometimes	quite often	very often
(a) given rules and explanations in a way which focused attention directly on the formal features of language					
(b) provided rules and explanations for grammar and vocabulary incidentally, in order to facilitate the students' performance on some other, communicative, task					
(c) taught aspects of "paragraph and text" structure, e.g. adverbs such as first, next; clause and sentence or logical connectors (although, therefore, however) and other features relevant to the organization of ideas in a text					
(d) given exercises to students on grammar and/or vocabulary (i) orally					
(ii) in written form					

9. (continued)

	never	rarely	sometimes	quite often	very often
(e) taught the social and cultural rules of language, e.g. how to make polite requests, address strangers, express opinions, disagree politely, etc.					

CORRECTION

10. During oral class work, how did you deal with the errors made by your students?

	never	rarely	sometimes	quite often	very often
(a) I tried to correct errors immediately					
(b) I corrected only those errors which interfered with communication, and ignored the others					
(c) I tried to focus on the meaning of what the students said, and not so much on whether their use of language was accurate					
(d) I took note of the errors for later correction					
(e) Other (please specify below)					

11. When you corrected student's written work, which aspects of language and language use did you correct?

	never	rarely	sometimes	quite often	very often
(a) spelling					
(b) grammar					
(c) use of vocabulary					
(d) logical organization of text					
(e) stylistic appropriateness					
(f) other (please specify)					

12. When you corrected student's oral work, which aspects of language and language use did you correct?

	never	rarely	sometimes	quite often	very often
(a) pronunciation					
(b) grammar					
(c) use of vocabulary					
(d) stylistic appropriateness					
(f) other (please specify)					

EVALUATION

13. What aspects of language and language use formed the focus in the tests you gave to your grade 11 students? We would like you to rank order the items below. If there are any other items that you included, please note them down underneath the list provided and then rank order the whole list.

Assign a value of "1" to the most important item, "2" to the next most important, etc. You may assign the same rank to more than one item if you feel they are of equal importance.

	oral tests	written tests
(a) accurate use of grammar		
(b) accurate use of vocabulary		
(c) accurate spelling		
(d) accurate pronuncation		
(e) stylistic appropriateness		
(f) logical organization of text		
(g) ability to get one's meaning across (irrespective of linguistic or stylistic errors)		

HOMEWORK

14. Since September 1984, how often have you assigned the following types of homework to your grade 11 students?

	never	rarely	some-times	quite often	very often
(a) creative writing tasks, e.g. compositions, letters, reports, projects, etc. with no models or guidance provided					
(b) memorization of dialogues, poems, etc. for oral presentation					

14. (continued)

	never	rarely	some-times	quite often	very often
(c) reading extended text, i.e. longer than half a printed page					
(d) grammar and vocabulary exercises in single-sentence format					
(e) grammar and vocabulary exercises which require extended answers (i.e. two sentences or more)					
(f) creative oral tasks: student-prepared oral presentations, such as role-play situations, oral reports, etc.					
(g) guided writing tasks based on pictures, diagrams, picture sequences, fill-in-the-blank model paragraphs, etc.					
(h) comprehension questions (or other types of questions) which require writing two or more sentences					
(i) preparation of scrapbooks, maps, collages, or other combinations of printed and visual materials based on topics of personal interest to the students					

(j) Please add examples of any other type of homework you have frequently assigned.

USE OF FRENCH

15. How often did you use French with your students for any interactions occurring before and after the actual French period?

never	rarely	sometimes	quite often	very often
( )	( )	( )	( )	( )

16. Did you encourage the students to use French during the French period for purposes of routine administration and classroom management?

never	rarely	sometimes	quite often	very often
( )	( )	( )	( )	( )

17. How often did your students try to use French before and after the actual French period?

never	rarely	sometimes	quite often	very often
( )	( )	( )	( )	( )

Thank you very much for your co-operation. Please return the questionnaire to Françoise Pelletier, Modern Language Centre, OISE, 252 Bloor Street West, Toronto M5S 1V6.

APPENDIX F

HOOL: GRADE(S): DATE: VISIT NO.: Part A

ACHER: LESSON (Min.): OBSERVER: Page

BJECT:

1.1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	ACTIVITIES & EPISODES	PARTIC. ORGANIZATION							CONTENT																				CONTENT CONTROL			STUDENT MODALITY					MATERIALS							
		Class			Group		Ind.		MAN.		LANGUAGE				OTHER TOPICS																						Type				Source			
															NARROW				LIMITED					BROAD													Text							
		T ↔ s/c	S ↔ s/c	Choral	Same	Different	Same	Different	Procedure	Discipline	Form	Function	Discourse	Soc. Ling.	Classroom	Formul.Ex.	Pers./Bio.	Other	Personal	Rout./Soc.	Fam./Com.	School T.	Other	Abstract	Pers./Ref	Imagination	World T.	Other	Teacher/Text	T./Text/Stud	Student	Listening	Speaking	Reading	Writing	Other	minimal	extended	Audio	Visual	L2	L1	L1-adapted	Stud.made

Allen, J.P.B., Fröhlich, [illegible] and Spada, N. COLT

hool:
ver:
ect:
e of visit:
:

TEACHER VERBAL INTERACTION

STUDENT VERBAL INTERACTION

Part B

	COMMUNIC. FEATURES			TARGET LANG.		INFORMATION GAP				SUST. SPEECH		REACTION CO/MES.		INCORPORATION of S. UTTERANCES								TARGET LANG.			INFORMATION GAP				SUST. SPEECH			FORM RESTR.		REACTION CO/MES.		INCORPORATION of S/T UTTERANCES						
						Giving Info.		Request Info.																	Giving Info.		Request Info.															
			Off-task	L1	L2	Predict	Unpred.	Display Request	Info. Request	Minimal	Sustained	Code	Message	Correction	Repetition	Paraphrase	Comment	Expansion	Clarif. Request	Elabor. Requ.	OFF-task	L1	L2	Disc.-Initiation	Pred.	Unpred.	Display Request	Info. Request	Ultraminimal	Minimal	Sustained	Restricted	Unrestricted	Code	Message	Correction	Repetition	Paraphrase	Comment	Expansion	Clarif. Request	Elabor. Request
A	E	M	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40

COLT Observation Scheme: Revised definition of Categories

The COLT observation scheme is divided into two parts. Part A describes classroom events at the level of episode and activity, and Part B analyses the communicative features of verbal exchanges between teachers and students or students themselves as they occur within each activity.

Part A: Classroom events

I. Activity

The first parameter is open-ended, that is, no predetermined descriptors have to be checked off by the observer. Each activity and its constituent episodes are separately described: e.g. drill, translation, discussion, game, etc. (separate activities); teacher introduces dialogue, teacher reads dialogue aloud, students repeat dialogue parts after teacher (i.e., three episodes of one activity).

II. Participant organization

This parameter describes three basic patterns of organization:

1. Whole class

a) Teacher to student or class, and vice versa (One central activity led by the teacher is going on; the teacher interacts with the whole class and/or with individual students).

b) Student to student, or student to class and vice versa (Students talk to each other, either as part of the lesson or as informal socializing; one central activity led by a student may be going on, e.g. a group of students act out a skit and the rest of the class is the audience).

c) Choral work by students (The whole class or groups participate in the choral work, repeating a model provided by the textbook or teacher).

2. Group work

a) Groups all work on the same task.

b) Groups work on different tasks.

3. Group and individual work

a) Individual seat work (Students work on their own, all on the same task or on different tasks).

b) Group/individual work (Some students are involved in group work, others work on their own).

III. Content

This parameter describes the subject matter of activities, that is, what the teacher and the students are talking, reading or writing about or what they are listening to. Three major content areas have been differentiated, plus the category 'topic control':

1. Management

Procedural directives and disciplinary statements.

2. Explicit focus on language

a) Form: Explicit focus on grammar, vocabulary, or pronunciation.

b) Function: Explicit focus on illocutionary acts such as requesting, apologizing, and explaining.

c) Discourse: Explicit focus on the way sentences combine into cohesive and coherent sequences.

d) Sociolinguistics: Explicit focus on the features of utterances which make them appropriate to particular contexts.

3. Other topics

This is a tripartite system which deals with the subject-matter of classroom discourse, apart from management and explicit focus on language.

a) Narrow range of reference

Topics of narrow range refer to the immediate classroom environment, and to stereotyped exchanges such as 'Good Morning' or 'How are you?' which have phatic value but little conceptual content. Included in this category are routine classroom references to the date, day of the week, weather, etc.

b) Limited range of reference

Topics of limited range refer to information which goes beyond the classroom while remaining conceptually limited: movies, hobbies, holidays, school topics including extracurricular activities, and topics which relate to the students' immediate personal and family affairs, e.g. place of residence, number of brothers and sisters, etc.

c) Broad range of reference

Topics of broad range go well beyond the classroom and immediate environment, and involve reference to controversial public issues, world events, abstract ideas, reflective personal information, and other subject matter, such as math or geography.

Topic control

Who selects the topic that is being talked about, the teacher, the student or both? If the teacher selects the topic, this may be done in conjunction with a textbook.

IV. Student modality

This section identifies the various skills which may be involved in a classroom activity. The focus is on the students, and the purpose is to discover whether they are listening, speaking, reading or writing, or whether these skills are occurring in combination. A category 'other' is included to cover such activities as drawing, modelling, acting, or arranging classroom displays.

V. Materials

This parameter introduces categories to describe the materials used in connection with classroom activities.

1. Type of materials

a) Text (written).

b) Audio.

c) Visual.

2. Length of text

a) Minimal (e.g., captions, isolated sentences, word lists).

b) Extended (e.g., stories, dialogues, connected paragraphs).

3. Source/purpose of materials

a) L2 (specifically designed for L2 teaching).

b) L1 (materials originally intended for L1 or non-school purposes).

c) L1-adapted (utilizing L1 materials or real-life objects and texts, but in a modified form).

d) Student made (materials produced by the students themselves).

Part B: Communicative features

Seven communicative features have been isolated:

I. Use of target language

(a) Use of first language (L1)

(b) Use of second language (L2)

II. Information gap

This feature refers to the extent to which the information requested and/or exchanged is unpredictable, i.e., not known in advance. The two categories designed to capture this feature are:

1. Requesting information

(a) Display request (The speaker already possesses the information requested).

(b) Information request (The information requested is not known in advance).

2. Giving information

(a) Relatively predictable (The message is easily anticipated in that there is a very limited range of information that can be given. In the case of responses, only one answer is possible semantically, although there may be different correct grammatical realizations).

(b) Relatively unpredictable (The message is not easily anticipated in that there is a wide range of information that can be given. If a number of responses are possible, they can provide different information).

III. Sustained speech

This feature is intended to measure the extent to which speakers engage in extended discourse, or restrict their utterances to a minimal length of one sentence, clause or word. The categories designed to measure this feature are:

1. Ultraminimal (utterances which consist of one word - coded for student speech only).
2. Minimal (utterances which consist of one clause or sentence - for the teacher, one-word utterances are coded as minimal).
3. Sustained speech (utterances which are longer than one sentence, or which consist of at least two main clauses).

IV. Reaction to code or message

Explicit code reaction (a correction or other explicit statement which draws attention to the linguistic incorrectness of an utterance).

Explicit message reaction (a correction or other explicit statement which draws attention to the factual incorrectness of an utterance).

V. Incorporation of preceding utterances

To allow coding for a limited selection of reactions to preceding utterances six categories have been established.

1. Correction: Correction of previous utterance/s.
2. Repetition: Full or partial repetition of previous utterance/s.
3. Paraphrase: Completion and/or reformulation of previous utterance/s.
4. Comment: Positive or negative comment (not correction) on previous utterance/s.
5. Expansion: Extension of the content of preceding utterance/s through the addition of related information.
6. Clarification request: Request for clarification of preceding utterance/s.
7. Elaboration request: Request for further information related to the subject-matter of the preceding utterance/s.

VI. Discourse initiation

This feature measures the frequency of self-intiated turns (spontaneously initiated talk) by students.

VII. Relative restriction of linguistic form

Two categories have been proposed to examine the degree of restriction placed upon student talk:

1. Restricted: The production or manipulation of one specific form is expected, as in a transformation or substitution drill.
2. Unrestricted: There is no expectation of any particular linguistic form, as in free conversation, oral reports, or personal diary writing.

Chapter 5

THE IMMERSION OBSERVATION STUDY

Merrill Swain and Susanne Carroll

1. INTRODUCTION

As a result of the grade 6 immersion proficiency study (Year 2 Report), and of considerable other data that have been collected over the years in the Modern Language Centre concerning the interlanguage used by the immersion students (e.g. Harley 1986; Harley and Swain 1978, 1984, Swain and Lapkin 1982), it became increasingly obvious that to better understand the proficiency results obtained, and develop materials to improve them, what was needed was an accurate description of what actually occurs in immersion classes. We needed to understand the process that led to the second language proficiency 'product' we had been observing. Therefore an observation study was undertaken.

The observation study was carried out in nine grade 3 and ten grade 6 classes in three Ontario school boards. Two of the school boards were located within the Metropolitan Toronto area and one was from the Ottawa area. In one board, five grade 6 classes and four grade 3 classes were observed. In another board, five grade 3 classes were observed. And in the third board, five grade 6 classes were observed. Thus, in all, ten grade 6 classes from two boards and nine grade 3 classes from two boards were observed. The observations were carried out during the months of April and May, 1984.

The basis for the selection of classes within boards was as follows. In the two Metropolitan Toronto boards, the selection of classes was determined by the total number of available immersion classes at specific grade levels. Thus in the board where both grade 3 and grade 6 classes were observed, there was, at the time of the data collection, a total of five schools offering a daily French immersion program at the grade 6 level. For purposes of comparability, it was decided to observe five grade 6 and five grade 3 classes in the same schools; however, one grade 3 teacher preferred not to take part in the study, hence the lower total number of grade 3 classes in the study. In order to make up a larger total sample at each of the grade levels, further grade 3 and grade 6 classes were sought in the Metropolitan Toronto area. In one board, five grade 3 classes could be found and were incorporated in the study, but there were insufficient numbers of grade 6 classes. It was decided therefore to extend the grade 6 sample to the Ottawa/Carleton area, where a representative sample of five grade 6 immersion classes was selected with the help of board personnel.

The observations were conducted over the full school day. In the grade 6 classes this involved both an English and French instructional period. In the grade 3 classes, most of the classes were conducted completely in French. A stereo audiocassette recorder with two microphones ("Sound Grabbers") attached to two walls was used to capture the public talk of teachers and students.

An observer was present in the classroom while the taping was being done. The observer used the COLT observation scheme (see Chapters 3 - 4) in precisely the manner used in the Core French Study (see Chapter 4 for details). The observer also changed the tapes as required, and noted text material in common use in the classrooms.

After the observations, the classroom talk (comprising student and teacher talk) was transcribed using ordinary French orthography.[1] The transcriptions were done by a fluently bilingual native speaker of French. The grade 3 transcripts were checked by a second fluently bilingual native speaker of French. The grade 6 transcripts were corrected by a fluently bilingual native speaker of English. The type of teamwork based on the linguistic competence of the researcher which was used in checking the grade 6 transcripts proved useful: some of what was undecipherable to one was decipherable to the other as teachers were usually native speakers of French, while students often had a production influenced by their knowledge of English.

The data collected are exceedingly rich. What is reported in this chapter and in Appendix A, pp. 264-286, represents analyses of only some aspects of language use in immersion classes. Other analyses will be carried out as time and finances permit. In the present chapter, we report on the following analyses:

Vocabulary instruction in immersion classes. In this section a description of how vocabulary is taught in grade 6 immersion classes is provided. The description provides insight into the results in Chapter 1, "Perspectives on Lexical Proficiency in a Second Language".

Vous/tu input. In a paper "Aspects of the Sociolinguistic Performance of Early and Late Immersion Students", Swain and Lapkin (in press) report on immersion students' sociolinguistic use of tu and vous. In order to explain their finding that early immersion students overuse tu in formal situations, the uses of tu and vous by grade 6 teachers and students are examined.

Student talk in teacher-fronted activities. Opportunities to use a second language are considered important in developing learners' grammatical, discourse and sociolinguistic competence. For this reason, the frequency and length of (transcribable) student talk in both the French and English portions of the grade 6 classes are examined. Additionally, the sources of student talk in French are investigated in order to determine if any one particular source encourages the use of extended turns.

Error treatment. In this section, the frequency and type of surface structure errors grade 6 students make in their spoken French, and the extent to which teachers correct them, are investigated. In this context, the consistency and persistency of feedback concerning learner errors provided by teachers are considered.

Although one goal of immersion is to learn language through learning content, a general observation about the classes is that form and function are kept surprisingly distinct. They are kept isolated from one another in two ways.

First, there is usually a time set aside for grammar, and during that time formal rules, paradigms and grammatical categories are learned. It is relatively rare that teachers refer to what has been learned in a grammar lesson when involved in other subject related lessons, or that content based activities are set up specifically to focus on form related learning.

Second, the forms of the categories they are learning are separated from their meanings. In the example below, the student is confused because the same form "les" has two completely different functions. The form dominates his thinking, rather than the meaning the form conveys. The students in this grade 6 class (No. 156) have had a lesson on personal pronouns and are involved in completing a written exercise as seat work.

S Est-ce que...comment...est-ce que uhm "les" est un pronom personnel?
T Regarde dans ton tableau.
S Monsieur.. un pronom, c'est l'article?
T C'est/c'est un article quand il est placé devant un nom. Si/si il prend la place d'un nom, ce n'est plus un article. Regarde ici, tu l'as.
S Je sais.
T La L'exemple. C'est toi-même qui l'as lu. "Maman appelle Luc et Jean. Elle les/les/elle LES attend."
S Je sais.
T "Les" prend la place de qui?
S Oh. Les euh uhm les enfants.
T Voila
S Oui mais X X ça dit LES avant et ça dit pas ou X X X X
T À ce moment-là, il n'est pas/il n'est pas pronom personnel..."les élèves".

In immersion classes, where language and content learning are equally important goals, a closer alliance than has been observed between the teaching of structure and meaning deserves our future attention.

2. VOCABULARY INSTRUCTION IN IMMERSION CLASSES

2:1 Introduction

There has been considerable focus in recent years in second language studies on the acquisition of various aspects of grammatical knowledge, for example acquisition of syntax, phonology and semantics (as a perusal of relevant journals will attest). Studies on the acquisition of second language vocabulary are much rarer,[2] so much so that one prominent writer in the field has referred to the situation as one of "neglect" (Meara 1981, see also Meara 1984, in press). If we know very little about how second language learners in various kinds of environments learn what words mean, how they are constructed, and how they are used in particular speech situations (but see Abberton 1968, Blum-Kulka and Levenston 1983, Gass to appear, Levenston and Blum 1977, Yoshida 1978, and Chapter 1), we know even less about how teachers teach that information (but see West 1954). It would be simplistic, and wrong, to view vocabulary teaching as merely the institutional complement of vocabulary learning (since what gets taught is not necessarily what gets learned), nonetheless it would be unreasonable to suppose that the structure and type of activities in the classroom environment have no effect on the nature of the learners' developing L2 lexis. If the classroom environment is the sole or even the major source of input for learning the L2, then its influence for determining what gets acquired would appear to be a necessary one.

Having said this, we have of course said nothing about the psychological processes inherent in vocabulary learning, nor how they are invoked or utilized. We have said nothing about how the environment might trigger or prevent learning but it would seem to follow from the remarks above that there is a utility to and desirability of studying the instructional context, to see how teachers interpret the task of vocabulary instruction, to see what kind of input is available for certain kinds of learning, and to see how constraints on teacher-student or student-student interactions might highlight or hinder specific aspects of lexical learning. It is known that the acquisition of individual lexical items is incremental, potentially limitless and heavily constrained by the individual's experience, characteristics which appear to distinguish it from other kinds of language learning. It may be the case, therefore, that lexical acquisition in the second language context could be explained in terms of a theory of learning (understood as a theory of hypothesis testing unconstrained by specifically linguistic parameters) where the language user's proficiency develops out of a complex interplay between existing forms of knowledge and the specific linguistic environments in which the hypotheses are formulated and tested. If such speculation has a priori plausibility, it becomes imperative to have descriptions of the relevant learning environments.[3]

2:2 Objectives of the Study

It is the purpose of this study to provide a modest beginning to the investigation of the pedagogical issues related to vocabulary instruction in second language contexts and to provide a descriptive categorization of vocabulary-related classroom activities which may, we hope, be relatable to studies of lexical knowledge and use where subjects learn what they know about the words of their L2 in the kinds of environments described here. The investigation does not include a study of either linguistic competence or proficiency as they relate to lexis, and one would like to see studies which explicitly link what gets done in the classroom with the knowledge base that students acquire and the capacities they develop for exploiting it in various contexts. The discussion will, however, bear on how vocabulary appears to be selected and taught in various grade 6 French immersion classrooms (which will be described in the next section). The categorization will be couched in terms of the grammatical, discourse and sociolinguistic competences discussed in Canale and Swain (1980) and the Year 2 Report. In other words, observed classroom behaviour will be analysed with a view to determining the extent to which it could potentially contribute to the development of these competences. The discussion will be non-quantitative[4] and the conclusions of a suggestive nature. The tentativeness of the results is necessitated at least in part by the dearth of theoretically relevant studies. These remarks notwithstanding we shall spell out what sorts of things are to be acquired when one learns a language's lexis fully or proficiently, thereby making explicit the kinds of information that teaching must attempt to get across either explicitly (through explanation, exercises, practice and so on) or implicitly (through exemplification in speech and written discourse). We shall also attempt to make explicit where activities are likely to fall short of developing mastery of the relevant types of knowledge.

The report of the study is organized as follows: Section 2:3 provides a brief description of the classes observed, the methodology of the observations and the analysis, and the types of texts and teaching materials that were examined. Section 2:4 provides a set of instructional foci, cutting across a number of classificatory labels. These labels provide ways of looking at instructional activity. Section 2:5 provides a set of descriptors for language learning. Here we make some very general remarks about learning and attempt to provide a certain background to discussions of learning and the environment. Section 2:6 gives a typology of various types of information that are revealed in and via words (and larger lexical expressions) as well as a discussion of the general organization of the lexicon. This typology is structured on the basis of cognitive descriptors (provided by linguistic theory) rather than pedagogical ones. Section 2:7 follows with a discussion of how the various pedagogical types could (at least in principle) facilitate the learning of the information discussed in section 2:6. Conclusions follow in section 2:8.

2:3 The Observations and Students

The study to be described here examined early French immersion grade 6 classes all of which (ten in number) came from two boards and the data discussed come from only the French-language segments of the transcriptions. The complete transcripts constitute the major source of data utilized.

Some texts and exercises used in the classroom were available and were also examined to see what kinds of written exercises and planned activities might be assigned which involve vocabulary instruction. A partial list of classroom texts and examples of reference materials appear in Appendix B, pp. 287-292.

2:4 A Classificatory Schema for Vocabulary Instructional Activities

In this section an analytical schema relating to activities for vocabulary instruction will be presented. It will consist of a descriptive classification of what can occur in the classroom along a number of different dimensions. These dimensions represent simply a number of distinct but potentially overlapping classificatory labels for aspects of pedagogical activities, teacher behaviour or methodology, techniques or other aspects of the teaching process (see Table 1, p. 252).

Planned/unplanned instruction

The first dimension to be examined is that of planned vs. unplanned vocabulary instruction by the teacher. This dimension identifies the extent to which vocabulary teaching is seen as a purposeful activity, with set objectives and subject to pedagogical constraints (arising from curricular, methodological and other teaching developments). It allows one to determine the place of vocabulary instruction with respect to other types of language-related instruction (grammar, spelling, writing, reading, etc.). If vocabulary instruction is a planned activity, it is likely to reflect current or accepted beliefs about how vocabulary fits into the larger scheme of language knowledge, what elements of lexical information should be taught (see below), what types or classes of words should be chosen, how the information should be presented to make it more readily learnable, etc. While it is not necessarily true that spontaneous instruction could not also be carried out in ways consonant with pedagogical objectives, it seems reasonable to suppose that spontaneous instruction will occur in response to immediate needs of the students and may or may not connect to the overall objectives of a given lesson. In other words, a child may want to know how to say "X" where the particular lexical item possesses none of the properties of items selected for instruction. One can then hypothesize that planned and unplanned vocabulary instruction will correspond to different situations, and different types of behaviour on the part of both teachers and students, and perhaps lead to different types of learning patterns. It may be the case, for example, that planning relates to e.g. which lexical items get selected for instruction and thus what items are brought to the students' attention. On the assumption that exposure and attention are necessary prerequisites to learning, then planning could strongly influence it. On the other hand, it might be the case that spontaneous instruction arises out of a real communicative need of the students, i.e. both gaps in knowledge and a desire to communicate specific information. It might be the case that such instruction correlates to higher motivation to learn. Clearly, all of these questions need to be investigated in a controlled fashion.

Systematic/haphazard instruction

A second and related dimension is that of systematic vs. haphazard instruction, or more precisely which aspects of lexical knowledge are taught in a systematic fashion and which are done haphazardly. Even if vocabulary instruction is recognized as an important part of language training and is included in the curriculum, this recognition does not mean that it will be carried out in a systematic fashion. The degree of systematicity of instruction of specific features can have an impact on whether they get learnt by the student and how well. So, for example, if one of the goals of vocabulary teaching is to get students to the point where they can spontaneously invent items to express certain concepts, then it stands to reason that one would want to systematically focus on rule-governed aspects of word-structure and transparent semantic patterns. To take a standard example from French, the semantic relation between adjectives and the corresponding adverbs of the form ADJ + ment is transparent and the rule of adverb formation is fairly straightforward. If instruction systematically requires the student to invent adverbs of the appropriate form, then one can anticipate greater automaticity and frequency in the actual production of adverbs of this form.[5]

The nature of the input: Written vs. oral language activities

A third distinction corresponds to the medium used for instruction. Many of the activities of the class are designed to teach knowledge and use of written language. Teaching literacy can be seen as one of the major functions of the school in our society and one can speculate that written language therefore takes on a privileged position inside the classroom not only for the encoding and transmission of information, what we might refer to as "propositional" language or (to use one of Halliday's terms) "ideational language" (see Gregory and Carroll 1978) in contrast to social, expressive or poetic language, but also as a symbol of what language is.[6] Thus, written language may be selected more often as the source of vocabulary, thereby restricting the range and types of input that learners get, and therefore what information is potentially available to be learnt in general, and possibly at particular stages of learning.

The distinction that we wish to draw here does not correspond to a simple difference in medium (or channel) of presentation. We are not simply claiming that sometimes we articulate words and sometimes we write them down. Rather the two media are typically associated with different varieties of language (on this notion see Gregory and Carroll 1978). Thus, it is well known that written language tends to be used to express more formal language.[7] This fact is reflected in the vocabulary used in each medium. Written language associated with formal registers can have a far more varied (sometimes a more precise) vocabulary than speech, and words which are acceptable in dialogue are often shunned in print.[8] Furthermore, the frequencies of specific words will vary according to the medium used, and we can assume that the frequency of occurrence of an item in the learners' environment must have some effect on its integration into their system. So words which are used more often ought to be easier to recall. Words which are heard more often or read more often ought to be recognized more readily.

Turning to considerations of a more grammatical nature, we can see that in some languages the actual structural nature of the words may vary depending upon the medium used and its conventions. It is well known that spoken idiomatic English contains a relatively higher percentage of derivationally simple words, which are monosyllabic, historically related to Old English forms and which do not illustrate certain segmental alternations and stress shifts associated with French and Latin-based words (c.f. Aronoff 1976, Chomsky and Halle 1968 on the structural and phonological properties of the "Latinate" and "Native" vocabulary) than do certain types of formal, written discourse (such as one finds in academic journals). In addition, there are certain derivational suffixes (e.g. -hood) which can only attach to roots that are of the so-called "Native" class. They do not attach to "Latinate" roots. Latinate words do typically occur in formal written English, which means that over-exposure to one medium could in principle affect the very type of grammatical learning the L2 learner can achieve. A learner heavily exposed to "street language" simply might not be exposed to sufficient exemplars of certain Latinate suffixes to learn specific rules of word formation. In contrast, someone who acquired English by reading only the writings of Conrad Black might not master the lexical sets associated with normal conversation. While the divisions in the French language vocabulary are not so dramatic,[9] they do exist. Greek-based prefixes in French have properties different from Latin-based prefixes and Greek-based words occur more often and in greater proportions in written formal French than in other varieties.

The correspondences between formal/informal language and written/spoken media are neither absolute nor necessary. Dialogue can be, and is, represented in many of the stories that children use in class. Nevertheless, it is difficult to represent styles of articulation, changes in rhythm and rate of articulation and other aspects of informal pronunciation in conventional orthography. Words which are frequently used but restricted in their distribution to certain speech styles, certain environments and/or one of the two media, words which are neologisms or have no conventional orthography may never occur in the texts that children see. If this is true, then L2 learners will have predictable gaps in their lexicons when input is highly dependent upon a given medium. Where such gaps are perceived, and when the aims of instruction include providing the means for conducting ordinary conversations on topics of everyday occurrence, then some means must be devised for providing the lexis which would be used for such conversations.[10]

Another way in which the two media tend to differ is in the explicitness of the expression of the speaker's intention. The written medium provides the possibility of communication independently of the original situation in which the text was produced. For writers' meaning and intentions to be clear, they cannot depend upon the readers' knowledge of that situation in order to make the appropriate inferences. Writers must put into words what it is they want to convey and they must draw connections for the reader since there will be no occasions for questioning the writer about what he or she meant. This is clearly true to a much lesser extent of spoken language. One can be vague, misleading or even confused on occasion because one can usually make repairs.

At the same time that one needs to be more exact and to the point in the written medium, one also needs to be more concise. Karmiloff-Smith (1981) has documented very clearly the extent to which child speech is highly redundant in that children will spread out over several propositions the information that an adult would convey within a single noun phrase. Spoken language is more often associated with this kind of redundancy. Although it would be incorrect to suggest that there are no temporal (or length) constraints on speech, nonetheless the demands of "good style" are not the same.

Children are expected to acquire some understanding of these constraints and one can find activities designed to teach them to respect the rule: Do not write the way you talk (see, for example, Genouvrier and Poulin 1973).

Pointing the way: Appealing to prior knowledge of the L1 vs. appealing directly to the L2 data

The distinction between spoken and written language raises the issue (at least for French and English speakers) of the extent to which teaching relies on and emphasizes prior knowledge, in particular resemblances between the L1 and the L2, vs. teaching on the basis of information inherent in the L2 alone. In their written forms, English and French often look more alike than they sound alike. This fact coupled with the special role of the written media in the classroom means that teachers can consciously encourage students to develop cross-language associations, building on properties associated with specific items in English to direct the learning process of specific words in French. In contrast to this type of approach is one where the teacher relates particular bits of data to other forms of the L2 without appealing to knowledge of the L1 or even discouraging any direct linking of potentially relatable items. There is a common perception among students and teachers alike that frequently resorting to the L1 in the L2 classroom is generally speaking a bad thing, the idea being that one cannot learn to communicate "in the L2" if one is communicating following patterns established in the L1. There is a natural logic to this argument but one should not let it obscure the fact that controlled use of the L1 for well-defined purposes could conceivably enhance certain kinds of learning; the problem arises in determining which kinds of learning will be enhanced and which kinds will not.

At the moment, it appears that neither psychology nor linguistics has a great deal to say on this issue which can be considered as reliable and sure. But it is generally agreed that learning occurs by using in particular ways what one already knows and connecting it to, e.g., perceptual events (see Anderson 1985) and it seems reasonable that second language teaching should exploit this fact. Included in what the L2 learner already knows is not only a certain amount of information about the L2 but also vast amounts of information about the L1 and consequently about language and communication more generally. It might be useful to consciously exploit such knowledge rather than ignoring it or taking it for granted. If one could make precise if and how resorting to knowledge of English facilitates learning certain aspects of French (and only research on learning will do this, as an example see Hammar 1978), then one could build in teaching techniques or strategies which would use this knowledge. At the same time one could eliminate or attempt to avoid occasions where resorting to English actually hinders acquiring properties of French.

It would appear that written and oral media might differ in how they can be utilized for either appealing to cross-linguistic resemblances or appealing directly to the data. There are similarities in the graphic representation of English and French words which appear to be greater, and hence more salient, than resemblances in pronunciation. If one therefore chose to incorporate into one's teaching planned exercises focusing on such things as homonyms, the identification of cognates, word expansion games, and so on, and one wanted to exploit the learners' knowledge of the L1 in order to structure the acquisition of the L2 lexis, then one would want to resort to the medium where the similarities are more readily perceivable, namely writing. On the other hand, teachers who want to direct their students' attention to intralinguistic similarities might want to begin with oral forms or patterns.

Control of vocabulary selection

Another instructional dimension concerns the issue of who decides or controls what specific words will be taught, at what stage of learning, and in connection with what other features of the language. As reading lessons are often the source of vocabulary lessons, then the selection of words to be explicitly taught falls, to a large extent, under the control of the text author.[11] Furthermore, the choice of texts to be used may be made at the level of the board of education or even higher. The teachers' direct influence may only enter in when they decide to use one reading passage or one book rather than another (among the list of approved or available texts). Such decisions are often based on past classroom experiences with given materials and the teachers' understanding of what is "easy" or "hard" for their own students. The role that vocabulary selection plays in defining what counts as easy or hard has not been intensively studied from a theoretical point of view (but see Chau 1972, Higa 1965, and Rodgers 1969), however it seems reasonable to assume that a lot of unfamiliar vocabulary in a reading passage will render its meaning more obscure. It is therefore possible that the degree of unfamiliarity of words from the teacher's perspective will play some role in deciding what passages are read, and what words are chosen for commentary, explanation and intensive study. One can note in passing that the obscurity of the text will be heightened if the unfamiliar items do not resemble previously known words (thus permitting certain kinds of guessing as to what the new words might mean) or if the topic of the reading is also unfamiliar (and hence does not permit guessing on the basis of prior knowledge of the subject matter). One might then hypothesize that effective vocabulary expansion techniques might involve using readings on highly familiar topics which contain unfamiliar words, or alternatively using readings which contain unfamiliar words which are derivative of words already seen.

One could also predict that vocabulary selection in more communicative classrooms will be decided indirectly on the basis of what children already know in the second language and what they can be assumed to understand in their mother tongue, both of these variables reflecting their experiences and their interests. It is one of the objectives of communicative approaches that instruction revolve in part around the learner's needs.[12] Text writers may only have general notions about what a student at a particular grade level ought to understand.[13]

In contrast to the above situation, vocabulary selection is very much under the teacher's full control in oral exchanges. And there is considerable evidence to suggest that possibly universal principles of linguistic interaction guide linguistic performance between (speech) actors who do not share the same degree of information. Adults (and older or linguistically more competent children) modify their production patterns according to their perceptions of the child's ability to understand (cf. Snow 1972a, 1972b, Snow et al. 1977). Perceptions of the difficulty of interpretation relate to information about the knowledge base of the listener (most five-year olds do not know a great deal about science and so explanations about the topic will be couched in familiar terms). Those perceptions will include assumptions about familiar and "hard" words. So one might predict that if teachers accommodate their vocabulary use to facilitate comprehension by their students then much of the vocabulary that the students will hear, they will already know. One can therefore anticipate a contrast between text-based lessons (where words may be largely unfamiliar) and oral exchanges (where words will be largely known). Again, only empirical study will reveal if this picture is correct.

Linguistic focus of teaching: Interpretation vs. structure

Independently of the medium used to teach vocabulary, and of the systematicity, generality or explicitness of instruction, one can distinguish the linguistic focus or orientation of the instruction. Is the purpose of the teacher to instruct about word interpretation, say during a text analysis? Does the teacher want to teach a word meaning?[14] Does the teacher want to teach structural information about the word in question or aspects of its pronunciation and spelling (see section 2:6 for details)?

2:5 Remarks on Aspects of Vocabulary Learning

In this section, we shall make some brief, and necessarily oversimplified remarks about aspects of vocabulary learning to provide a contrast with the teacher-focussed discussion presented above. We want to keep in mind, of course, that the real objectives of classroom activity are to affect what takes place in the learners' minds (see Table 2, p. 252).

Focussed vocabulary teaching and mental attention

One could relate the extent to which vocabulary instruction is planned to the extent to which learning can be focussed on a recognizable problem. It may or may not be true that instruction will appeal to conscious processes but it seems reasonable to assert that in a lesson involving a series of activities whose explicit focus is vocabulary instruction the learners' attention will be focussed in some fashion on words. Thus, mental attention can be utilized for learning something about those words. This situation is to be contrasted with one in which the learners' attention is directed to some task other than that of learning about vocabulary but where words are involved and in which the teacher somehow expects that learning will occur. So vocabulary learning exercises will be effective only under certain conditions, namely when other aspects of the tasks are not so difficult as to use up the available (but limited) attentional resources (see Anderson 1985). It would appear as if the old adage "One thing at a time" can offer some pertinent advice to teachers. If the objective of a lesson is to learn some kind of information about new words (their spelling, pronunciation, their meaning, their grammatical properties, etc.), it would appear to be true that less learning will occur if students are not only presented with new items but also new activities as well. What attention might be directed to the task of learning the words, will be diverted to learning the parameters of the activity instead. It stands to reason, therefore, that vocabulary learning might be greater if new information about words or new items themselves are introduced during highly routinized activities (all other things — including motivation — being equal).[15]

Systematic vs. idiosyncratic learning

Another learning dimension involves systematic vs. idiosyncratic learning, that is to say the extent to which it involves some information unconnected to anything else. One must be careful to distinguish at least three issues here. The first involves the fact that word-learning is to some extent incremental, as noted above. There is a considerable amount of word-learning which involves specific, ungeneralizable information (e.g. that *chien* is pronounced [ʃjɛ̃], that it means *dog*, that *barder* is pronounced [baʁde], that it is a verb, that it means that someone is making a scene, that its subject is obligatorily *ça*, etc.). So these aspects of word-learning are not part of any system which can be used to predict their presence.[16] Since idiosyncratic facts about words are not predictable, it must be true that exposure to the specific lexical item is required to learn these things. Idiosyncratic aspects of vocabulary knowledge contrast with generalizable features (such as the information that the suffix *-ment* attaches to adjectives). One need not have to wait to hear a given adverb in *-ment* to "know" it, or at least predict it as a possibility in the language. One can invent the word on the basis of one's knowledge of adverb formation rules. To use an example from English, any native-speaker can acquire the noun *punk* and proceed to derive without prior exposure both *punkish* and *punkishly* on the basis of the knowledge that *-ish* is an adjective-forming suffix which attaches to nouns, and that *-ly* is an adverb-forming suffix which attaches to adjectives. So the first distinction arising from systematic vs. idiosyncratic vocabulary learning reflects the extent to which various properties of words must be learnt by direct exposure to the word or can be inferred from knowledge of other aspects of the grammar. This distinction will have an obvious bearing on which items are selected for explicit presentation, and also what aspects of vocabulary knowledge it may necessary to present in an explicit fashion through operations and manipulations.

The second issue involves the extent to which words are associated with a specific semantic domain. If learners are taught about a given topic, say parliamentary democracy, then there are specific sets of words which will be learnt because they denote objects, events, processes and characteristics associated with that topic. As humans, we organize our realities into systems and subsystems and our language must reflect this. It is not surprising that verbs describing eating and drinking events and activities appear with nouns referring to foods or edible and potable objects. The linguistic system merely reflects the system in the real world. So learning is systematic or unsystematic to the extent that words collocate or cluster in language expressing certain themes or semantic domains.[17] Associative games can lead to the explicit linking of pairs like *black/white*, *président/république*, *apple/paradise*. The systematicity here arises in part from the way in which the non-linguistic world is structured, and from the fact that in certain kinds of contexts, the collocations or frequency of co-occurrence of such pairs is psychologically salient.

The third issue arising out of the systematic/unsystematic dimension relates simply to the extent to which the linguistic input together with language instruction is coherent. For example, one learns passive participles during lessons on the passive, and all passive participles share certain properties. Consequently, one can expect lessons on the passive to facilitate the learning of passive particles. To the extent that the content of teaching is systematically chosen with an eye to linguistically relevant generalizations, and the properties available in the data are perceivable, then systematic learning ought to be possible (all other things being equal) if the teaching and the data permit appropriate cognitive processes to operate. Where the rules and the data are not consistent, or when they are simply not pertinent for the appropriate cognitive systems, then we can anticipate the predicted learning will not occur. The literature on conscious versus incidental learning can shed some light on these points. It suggests that an intention to learn is in fact not required for learning to occur. It is patently obvious that language learning can occur without there necessarily being explicit or conscious desire to learn a given rule or principle. Such is the case with first and second language learning in natural contexts. Conversely, good intentions will not overcome insufficient data or inadequate rules. Teachers should keep in mind, therefore, that learning can occur during interactions between the students and the teacher, the students among themselves, contacts outside the school etc. The teacher may have little or no control over such interactions and can anticipate that the students' own personal experiences may enhance, be neutral with respect to, or even conflict with explicit instruction in the classroom. Worse, the teachers' own verbal behaviour, i.e. the linguistic models that they present to their students, may not be consistent with the rules of the textbooks and grammars employed in the classroom. All of this may mean that the teacher will have little control over the necessary input for learning specific grammatical principles at a given moment in the student's learning process.

The effects of written vs. oral input on learning

We observed above (pp. 195-197) that input will vary according to whether language is encoded through speech or writing. It seems to be true that written symbols tap iconic memory and oral linguistic forms tap echoic memory (Anderson, 1985). It also appears to be true that writing and speech forms involve distinct cognitive systems. Evidence for this comes from the fact that brain injury or degenerative disease can affect one system without affecting the other. Not all aphasias are agraphias, and vice versa (cf. Albert and Obler 1978). The language user's system for recognizing written words may be disturbed so that he or she cannot indicate what that written word means, while at the same time being able to identify the same word encoded phonetically. Furthermore, to the extent that writing is not phonemic, it can lead to associations that the sound shape would not normally permit, thus triggering different kinds of learning behaviour depending upon how the language is encoded. This point is discussed more explicitly below.

Self- vs. teacher-initiated learning

Related to the issue of who selects vocabulary for instruction is the issue of who initiates the learning event. When teachers respond to requests for vocabulary from their students, it is the student who is determining which concepts will be expressed, and to some extent which forms will be rendered. Thus, when someone asks for a translation, it is often awkward to suggest anything other than an equivalent word, even if the translation is not appropriate (as we shall see later). And students do ask for words. On the other hand, when teachers control classroom activities, they are the ones who will decide what words will be learnt, and which concepts will be expressed by specific forms.

2:6 Knowledge In and Through Words

To aid the comparison of teaching practice and input with possible pedagogical goals, one should understand what one can possibly learn about language through its words. It should already be obvious that individual lexical items both belong to several distinct but interacting grammatical systems and relate language to the world by encoding meaning.

If it is true that words belong to grammatical systems, then it follows that the learning of the grammatical systems comes about to some extent by learning words. Vocabulary teaching could be organized in many ways to accomplish the teaching of these systems, the proper characterization of specific expressions, and suitable generalizations which would permit the genesis of other words. Teachers could present explicit and formal rules which encode the grammatical information. They can characterize the information in terms of simple rules of thumb which cover most cases but do not account for exceptions. They can simply present relevant linguistic data which instantiates generalizations and allow the students to induce a rule, and so on.

Regardless of the manner in which the information is couched, if the instruction is to be linguistically correct and real it must express the kinds of knowledge discussed below. It is this kind of knowledge that learners must acquire if they are to develop grammatical competence, discourse competence and sociolinguistic competence (see Table 3, p. 252).

Words as sound units

Words are realized through sounds and the sounds are organized into predictable patterns and abstract units (phonemes; onsets, rhymes and codas; syllables; feet; prosodic words; prosodic phrases. For discussion see Selkirk 1980). Since speakers of a language possess this kind of knowledge, one could consciously choose to teach it to the second language learner — either explicitly or indirectly through exemplification. It follows that vocabulary could be selected and presented in such a way as to illustrate the sound system of the language, and to encourage the correct recognition and production of the relevant sound units. This is a non-trivial task since it is well known that phonological rules and generalizations operate over various domains. Some low-level phonetic rules apply any time their context of application occurs (i.e. whenever the structural description of a formally-expressed rule is satisfied). This means that such rules could function within the syllable, the morpheme, the word, the phrase or the sentence, across sentences. Other rules only operate within a certain unit, as if the boundaries of that unit provided a wall beyond which the rule could not "see".[18] Consider as an example the rule of semivocalisation in French (Dell 1973). This rule states basically that an underlying high vowel changes to a semi-vowel in the context of another vowel. Examples occur in (1)

(1) a. écriviez [ekrivye] not *[ekrivie]
b. ciel [syɛl] not *[siɛl]
c. il y avait [ilyavɛ] [iliavɛ] (optional)

In the above examples, a word-internal combination of /i/ and Vowel leads automatically to a change in the form of the /i/. Similarly, when the operation of suffixation produces a sequence of root (écrive) + i (the imperfective suffix) + ez (the second person plural suffix), it creates a context where the rule can apply. However, there are contexts where the rule does not apply, namely when the high vowel is in one word and the conditioning vowel is in another word.

(2) a. joli ami [ʒɔ liami] not *[ʒɔlyami]
b. une fantasie ironique [ünfãtaziironik]
not *[ünfãtazyironik}

Neither type of phonological entity would present problems if learners merely repeated the sounds offered to them as appropriate pronunciations, i.e. if learning occurred only on the basis of imitation. Such a conclusion seems unwarranted, however. Both first and second language learners draw conclusions about the appropriate pronunciation for given spellings, and spellings can influence and interfere with the development of correct pronunciations.[19] One would therefore want to be sure that heavy use of written material was not having an adverse effect on the development of authentic pronunciations of words qua words, and, perhaps more importantly, of words as elements of discourse where written forms may not correspond at all to phonological words.[20]

Phonological rules that have restricted applications could be pedagogically interesting insofar as they reveal or mask morphological operations. Consider in this respect, rules like vowel nasalization. The exact status of this rule is controversial within the theory of grammar (cf. Dell 1973, Tranel 1982) but its theoretical description does not have any relevance to the pedagogical issue of drawing attention to the empirical facts, namely that vowels preceding /n/ are systematically nasal in some contexts and not others.

(3) a.	fin	[fɛ̃]	not *[fɛ̃n]
	fine	[fin]	
	finesse	[finɛs]	
b.	poupon	[pupɔ̃]	not *[pupɔ̃n]
	(Je) pouponne	[pupɔn]	not *[pupɔ̃], *[pupɔ̃n]
	pouponnière	[pupɔnyɛr]	not *[pupɔ̃nyɛr]
c.	divin	[divɛ̃]	not *[divɛ̃n]
	divine	[divin]	not *[divɛ̃]
	divinateur	[divinatœr]	not *[divɛ̃natœr],
	divinement	[divinmã]	not *[divɛ̃mã], *[divɛ̃nmã]
	divinité	[divinite]	not *[divɛ̃ite], *[divɛ̃nite]

Selecting and focussing on these sound rules would not only help to direct the learners' attention to the morphological relations between the selected items (more on this below) and thereby provide one necessary condition for learning to occur but would also permit the teacher to exploit spelling conventions to help structure their pronunciation. Thus, in the above examples, the spelling-pronunciation convention is "When a vowel precedes a word-final n the vowel is nasalized and the consonant is not pronounced, but whenever the nasal consonant is followed by a vowel or by another nasal consonant, then the vowel remains oral and the n is pronounced". Knowledge of these conventions can sometimes serve as a useful heuristic for analysing words for structure or meaning, as we shall see.

Words as morphological units

Traditionally, internal word structure has been analysed in terms of roots, stems, prefixes, suffixes and so on. These forms exist as structural entities independently both of the phonological units that realize them (the phones, phonemes, syllables, feet and phonological words) and the syntactic units that they themselves form.[21] Expressions like *maison* are morphologically simple consisting of an unanalyzable stem. They contrast with both *maisonnette* and *comparaison* which consist respectively of the stem *maison* + the suffix *ette*, and the prefix + root (= stem) *con + pare* + the suffix *aison*[22]. It is debatable to what extent there is any purpose in drawing students' attention to the internal structure of words whose affixation is no longer productive (which is true of *con + pare*), that is to say which could not be used to form new words, but there clearly is a point in teaching patterns of productive word-formation if one of the goals of second language instruction is to provide students with strategies for communicating. The point is that language learners should be encouraged to use what they know to expand their vocabulary (since they can never be taught all of the words of the language), and this strategy should include word-generation from language internal resources. Exercises which involve simplified word analysis, comparison of formal properties, and shared aspects of meaning could in principle facilitate the development of this ability by drawing the students' attention to the relevant generalizations. It goes without saying that the development of automatic use of such strategies in production (particularly oral production) will probably also require considerable practice.

In addition to knowing that words can have internal structure, native speakers know which expressions relate paradigmatically (either via word-formation rules, the productive cases, or via lexical redundancy rules, which provide an explicit association in the non-productive cases, see Jackendoff 1972)[23] to words belonging to other syntactic categories or to words of the same category which have related but somewhat different meanings. Native speakers know which cells of the table in (4) are empty.

(4) Word classes

Word Classes			
Verbs	Nouns	Adjectives	Adverbs
compare	comparaison	comparatif	comparativement
nourrit	*nourraison	*nourratif	*nourrativement
nourrit	nourriture	*nourritive	*nourritivement
nourriss +	nourrissons	nourrissant	*nourrissantement
joue	*jouaison	*jouatif	*jouativement
joue	*jouitture	*jouitif	*jouitivement
joue	joueur	jouable	*jouablement
---	jeu	*jeuatif	*jeuative
---	courage	courageux	courageusement
encourage	*encourage	*encourageux	*encourageusement
encourage	encouragement	encouragé	*encouragémment
encourage	---	encourageant	*encourageantement

Systematic instruction (of a planned or spontaneous type) on morphological structure should help students distinguish between existing words and words which are potentially available, on the one hand, and words which do not exist because they cannot exist. The expression jouage could be coined alongside of jeu but nourrisantement, jouantement are both impossible for grammatical reasons.[24]

Words as syntactic units

During the course of the preceding discussion, explicit mention was made of the syntactic status of words. All lexical items belong to some syntactic category and this fact determines where they may occur in the sentence. Furthermore, one finds associated with words other kinds of syntactic information. Some expressions can have plurality (nouns, determiners and adjectives) but only one category has inherent arbitrary singularity or plurality as a morphologically explicit property, namely nouns.[25] As an example, consider that both ciseaux and lunettes are obligatorily plural nouns (under a specific interpretation) referring to unique and perceptually circumscribed objects.[26] Individual nouns have gender. Knowing the word involves knowing its gender. Similarly, many word classes subdivide into special subsets whose individual members have distinct properties (such as conjugation type or declension type). Consequently, one can say that exercises which focus on the proper identification and classification of given words fall within vocabulary instruction too. Such classification hinges on first learning how to distinguish the sets (noun, verb, preposition etc.), then the various subsets (-er, -re or -ir verb type, for example) and finally the individual members. We do not know at this point how second language learners go about learning these distinctions, or indeed if they do it with any generality at all. Nonetheless, these and other syntactic properties (such as the special behaviour of expressions of quantity) could all be rendered salient through vocabulary selection and instruction.

Words as semantic units in discourse

Words are semantic units. They have meaning and in context they are assigned interpretations. Therefore, one can teach word meanings, or one can teach interpretations of expressions in specific contexts. Notice that while teaching learners about interpretation is an inherent part of communicative language teaching (since the semantics of expressions are "negotiated" in context), it does not follow that one need thereby spend class time on the meaning of expressions. This point will become relevant in discussion below where we see instances of specific context-embedded instruction about interpretation.

In addition to relating to events and objects in the real world, words will also be used to create text. As Carrell (1977) has pointed out, sequences of sentences by their coherence are seen to be cohesive. Grice (1967) has provided a theory to account for this fact. Moreover, some words have (or in some uses have only) the function of creating text, words such as toutefois, pourtant, néanmoins, typographical devices such as numbers or listings organized with premièrement, deuxièmement where the ordering is not temporal but logical, example:

> Premièrement, le pays a un déficit trop large.
> Deuxièmement, les revenus de taxes sont en baisse
> Donc . . .

Words as graphological entities

Words are graphological entities and, not surprisingly, learners spend a considerable amount of time learning the conventions regarding their orthography. This learning involves not only the spelling of individual lexemes but also general correspondences between sounds and spellings, regularities which are not replicated in sounds, comparisons with spellings in other languages. As we shall see below, exercises on homonymy, spelling and recognized problems such as agreement of homonyms in sentences are not ignored in immersion classrooms.

Words as sociolinguistic markers

Words belong to sociolinguistic systems and as such identify their users as members of particular communities. The characteristics words mark can be geographical, temporal, social, individual. The instantiations of sociolinguistic effects may implicate any of the grammatical systems (phonetic, morphological, syntactic, or semantic). Teachers ought to be aware that many words are not sociolinguistically neutral and they should decide to what extent they wish to go beyond vocabulary teaching of a highly general (user or situation independent) nature. In other words, to what extent does one provide information which allows the learner to express his or her identity, origins and attitudes? To what extent does one help learners adapt their behaviour to the attitudes and expectations of others? This type of instruction would entail making learners aware of distinctions drawn by other speakers. Vocabulary activities thus should make clear when words are restricted to one or two varieties and they should help provide detailed information about the kinds of situations in which certain words might appear and what kind of individuals might use them.[27]

2:7 Observed Classroom Activities

In this section some activities which were observed in the immersion classrooms will be described in detail. Analysis has revealed that only some of the possible categories of sections 2:5 and 2:6 were actually exploited to any extent. Some were not used at all.

Planned instruction: The reading lesson

The classroom observations confirm that most planned vocabulary teaching occurs during reading activities. Indeed, there is an even stronger claim to be made, namely that at the grade 6 level, the reading lesson is also a "vocabulary lesson". This is not to suggest that words were not taught at other times, clearly a false claim, but the activity most readily associated with planned and conscious teaching of vocabulary is reading. Observed activities included students reading aloud to ensure that they knew how to pronounce the written forms of the words, and to practise reading continuous text; questioning students as to story content; teachers and students providing explanations of the interpretation of specific passages; teachers providing explanations of words that they felt the students would not recognize.

These activities provided striking contrasts with teaching which was unplanned. The planned and conscious instruction tended to revolve around certain themes (e.g. air traffic controllers and their work, the slave trade, etc.) or certain speech functions (e.g.

how to convince, how to make logical arguments). In contrast, the "interactions" which gave rise to spontaneous instruction were often student-initiated, pedagogically undirected, unpredictable and irregular. The teachers perceived that students needed to know a word, where that perception could be characterized as "The student wants to know how to say ...". In other words, teachers would presuppose that the student wanted to express a certain concept or idea, and would provide a form for saying that.

In order to see the contrasts, compare (5) and (6)

(5)

Définissons les termes: Les centres urbains

1. Nommez les différents genres de communautés qui existent.
   Nommez en une pour chaque genre
   A) la campagne
   B) ____________________________
   C) ____________________________
   D) ____________________________
   E) ____________________________
   F) la super-cité

2. Qu'est-ce qu':
   A) une cité est: ____________________
   B) une super-cité: ____________________
   C) une ville tentaculaire: ____________________
   D) une mégalopole: ____________________
   E) une cité cosmopolite: ____________________

(6) No. 146
S1: Est-ce que je vas aller à onze heures à le bureau du...
du eum ' nurse '
T: de la garde

The objective of the exercises in (5) is clearly to fix a certain vocabulary connected to the theme of urban centres. They draw attention to distinctions in the meaning of the expressions within this one semantic domain. Systematic instruction therefore focusses on meaning. Presumably what is to be learnt are systematic semantic relations among the words. Note that the student can learn incidentally the gender of the vocabulary items, which words are compounds and which are not, that mégalopol is part of a noun but the sequence cosmopol is not, etc.

The dialogue in (6), in contrast, represents a kind of interruption of the normal pedagogic activities of the class. It is real communication, since the student has a need to find out a piece of information, and poses a question designed to get it. In the course of the student's speech act, the teacher observes, on the basis of a very common strategy for filling in gaps in lexical knowledge (namely resorting to English), that the student does not know the word for nurse and decides to supply it. Teachers often ignore this gambit and so learners cannot count on acquiring new words during conversations if they use it alone. The instruction here, then, is unplanned or spontaneous, any learning which occurs will be unfocussed in that it does not relate to the other activities of the class (or even to the student's desire to get permission to leave), it is incidental and teacher-initiated. Notice too that examples (5) and (6) contrast in that the planned exercise involves written language and the unplanned activity occurs in the oral medium. It would be safe to say that there is a general correlation between spontaneous teaching and the use of spoken French in the classes observed.[28]

Systematic learning and the reading lesson

It follows from what has just been said that systematic learning need not arise through the course of conversations in the classroom even if such conversations typify natural communication. Since the focus of conversation is the conveyance of information, speakers will focus on the content and not on form, on what they intend to say and not what they say. Not only will learners acquire individual lexical items independently of other words in the same semantic domain (hence there will be no explicit contrasts such as those seen in (5)), but the learning will be haphazard unless students learn to systematically ask for the words that they want to know. Acquiring such self-directed strategies must become a part of vocabulary-related instruction if it is to facilitate communication. Only then will students and teachers find a happy compromise between constantly interrupting the conversation (and frustrating the student's attempts to speak) and supplying useful and needed information.

Written vs. oral activities

It has already been noted that most planned vocabulary instruction in the classes observed arises in connection with reading. It follows that new words to be taught will be typical of written varieties of French. No attempts to teach words unique to the spoken mode were observed. Teachers did, however, use exercises and activities whose major purpose was to take the child's spontaneous, orally-rooted production and make it more like written French (cf. Genouvrier and Poulin 1973). Naturally, a great deal of effort was devoted to teaching spelling.

Spelling is, in particular, the focus of much of correction in writing. The comments of the teacher of class no. 116 show the extent of this emphasis

(7) No. 116
T: Euh nous corrigeons ensemble Alors écoutez bien...
l'orthographe compte évidemment Alors si c'est mal écrit c'est zéro

Spelling is the focus of planned teaching as well as of correction. From the same class

(8) No. 116
T: UN s'il vous plaît et vous l'épelez... Name!
S1: donne D-O-N-N-E
T: DEUX Name!
S2: porte P-O-R-T-E

Spelling is even the focus of some spontaneous teaching as in (9) below where the spelling out comes during a question-response correction sequence where spelling was not the point of the questions.

(9) No. 116
S1: ' bowling '
T: les quilles

S1: ' bowl '
T: Quilles Point cinq...les quilles (...) Q-U-I-L-L-E-S

This type of comment provides the student with a graphic form to associate with a given pronunciation and does so by involving the phonetic form of the orthography. This association is thus rather complicated since it implicates three different forms of the word (pronunciation, spelling-out-loud and a written form).

A simple association starts from the printed page

(10) No. 156
(student is looking at written exercises)
S1: Monsieur qu'est-ce que ça veut dire "éduqué"
T: Éduqué?
S1: Oui
T: Quel mot anglais qui te fait penser à "éduqué"
S: ' educate '

(later)
S2: Qu'est-ce que "éducté" veut dire? (student misreads word)
T: Éducté/éducté? (irrelevant text)
S2: éduqué
T: À quel mot est-ce que ça te fait penser?
S2: Mais j'ai/t'ai pensé c'était "éducté"

Using prior knowledge vs. first-hand learning

Observe that the purpose of the comments illustrated in (10) is to exploit the learners' ability to form associations with meaning on the basis of graphological form. Through such formal resemblances, the teacher can try to get at the meaning of the word without providing a translation directly. However, to the extent that the strategy relies on a knowledge of English, it does utilize translation rather than making a direct correspondence between the sounds of the word and a concept. Thus the strategy relies on prior knowledge (specifically of English spelling) and encourages the student to look for formal similarities between words of the language. If the appeal succeeds the learner will access any and all sundry information about the relevant English word whether or not it corresponds to identical information in French. It remains to be seen whether such blanket appeals facilitate learning all aspects of word learning (as opposed to e.g. recall and semantic recognition).

A somewhat different appeal occurs in (11). It constitutes an attempt to exploit spelling to access other French words.

(11) No. 326
T: Un bottillon de paille Qu'est-ce ça veut dire?
(silence)
Un bottillon de paille? D'où est-ce que ça vient ce mot-là "bottillon"
Regardez-le un peu
S1: bouteille
S2: beau
S3: bouteille

T: "botte" "botte" Je prends une botte, vous savez, une botte de paille
S3: C'est un mor/c'est comme de la paille qui est mis dans un morceau
T: Oui Vous vous rappelez le film/pas le film!/le ballet qu'on a vu ensemble qui était "La Fille Malgardée"? Vous vous rappelez les paysans qui sont venus avec des bottes de paille sur la tête/euh sur l'épaule?
S4 Et il y a un/
T: Quelque chose qui les attache ensemble
S5: Une corde
T: Corde Bon!

This example is indicative of how an astute teacher can draw freely on shared knowledge of the L2 to develop and render explicit structural knowledge and make analogies between forms the students have already mastered and a new word, even when this information was not the point of the lesson. The teacher begins by asking the students to focus on the word shape. The students respond by associating freely rather than providing a structural analysis. This in itself is interesting because it reveals that students are capable of sizing up gross similarities between spellings (notice that the sounds of the vowels in the three words are quite different) and accessing familiar vocabulary but have difficulty examining the word for internal structural analysis. Undaunted, the teacher provides the structural base, refers to past shared experience to provide a denotation for the word (more on this below) without resorting to the L1. Thus, the data suggests that the written word can be exploited both for making comparisons between L1 knowledge and the knowledge of the L2 to be acquired, and for making comparisons between prior knowledge of the L2 and the new information of the L2 to be learned. Although English and French lend themselves readily to cross-linguistic associations, teachers need not resort to this strategy as a first or only resort.

As a further issue, it was noted that although it would be perfectly possible to create exercises which focus on sound resemblances (rhymes, the number of syllables, the shape of the initial syllable, presence and location of stress), such exercises were absent from the observations. Indeed, the one category which seemed to be ignored was 2:6 (Words as sound units). Pronunciation was corrected systematically when students were reading aloud, and at random otherwise. No instances of activities focussing on the sound shape of words other than homonymy exercises (cf. Appendix B, pp. 287-291) and spelling activities were found. It would not be an exaggeration to claim that in the grade 6 immersion classrooms observed, word spellings but no other word form have primary, conscious attention. Both students and teachers seem to be more or less oblivious to the phonological form of words in planned activities and during text-based work. In short, what knowledge they possess about the sound shapes of words is being utilized solely on an unconscious level. While there is no evidence that not exploiting this knowledge in a planned way retards the students' language development, it is obvious that there is one kind of resource which is not being used at all.

A final comment is merited. Spelling, particularly when it involves agreement markers, will always constitute learning difficulties for immersion students (as for native speakers).[29] Teachers should therefore continue to emphasize correctness and to encourage students to develop self-help strategies such as the use of a dictionary. One might worry nonetheless about correctness becoming an obsession. The teachers that were observed seemed to exhibit a reasonable balance in their approach to the problem. There did not seem to be an overemphasis on spelling to the point where it interfered with communication or other instructional foci.

Focus on interpretation and meaning

An additional consequence of channeling vocabulary teaching through reading lessons was that there was a certain stress put on making meaning explicit. This is also a consequence of the heavy emphasis placed on meaning and interpretation. So, for example, students were assigned exercises where they had to provide a convincing argument as to why their own favourite television program was worth watching. Other students then rated the value of the arguments and the value was determined in terms of clarity of expression. Similarly, students were asked to assess the presentation of a descriptive (self-) written work read aloud to the entire class. The evaluation did not bear on the manner in which the text was read but rather on its clarity and coherence (where choice of vocabulary can play an important role). Spontaneous commentary from teachers also stressed the same thing:

(12) No. 146

(in a lesson on the slave trade, a student tries to explain why Europeans did not have sugar cane)

S1: exportent le rhum Oui ils ex/exportent euh/importent les choses à/aux/aux Antilles et ça va changer la vie parce que les Antillais sont.../parce que les Antilles sont euh/sont primitives

T: Primitives euh Je pense que je te suis...Tu as une bonne idée si tu peux le le clarifier

The teacher here was not satisfied with the vagueness of the word primitives and so insisted on clarification.

As noted previously, students regularly performed activities which required them to ask or answer questions related to text interpretation. They were also required to find synonyms, antonyms, or paraphrases for words or expressions in their readings, and to provide explanations for questions such as "Qu'est-ce que ça veut dire, X?" Students also were required to utilize the vocabulary being taught in exercises related to the text (e.g. in responding to questions), occasionally in unrelated activities such as writing exercises connected to general themes. It is unclear at this point to what extent students are given explicit instructions to use the words they have just learned in doing these writing activities (clearly a more controlled and goal-oriented procedure than simply allowing the students to use any vocabulary they choose). All of these types of activities involve teaching the meaning of words found in readings.

Meaning was also the focus of spontaneous exchanges in the class unrelated to reading. Students would often ask what individual words meant when they encountered them in French. Teachers also spent a certain amount of time verifying that their students understood them. Interestingly enough, teachers spent only minimal amounts of time asking students what they specifically intended in uttering a specific bit of language, and then only when the student was having obvious difficulty or when comprehension ceased. Such questioning rarely happens in the course of normal L1 communication because one can usually assume that speakers know their own language. It is only when speakers disagree about meaning that one might inquire "What do you mean when you say "X"? It would be a mistake, however, to conclude that the absence of overt confusion in the L2 classroom is proof that the students know that what they say conveys what they hope it will. This is simply because the absence of errors is not a sure indicator of real knowledge. There is an additional reason for not waiting for problems to arise to discuss issues of meaning and intention. Teachers spent a certain amount of time supplying vocabulary to students, in effect "putting words in their mouths". But it is obvious that at the same time that they are providing sound-shapes to convey concepts, they are also providing the concepts which are conventionally expressed by the word. The concepts may or may not be those students want to say since they may only have a very vague idea of what it is they want to communicate. Surely there is pedagogic value in systematically encouraging students to reflect on what they want to say and then choosing the vocabulary which will say precisely that. This problem goes beyond the question of clarity of description or argumentation for a defined speech act. Rather it is one of helping students learn that words have conventional meanings.

(13) No. 156

T: Huit: "Dans ce chapitre, quelle règle du protocole est de sa majesté?
euh protocole protocole protocole
S1: C'est quand il y a beaucoup de personnes commes les polices parce
qu'il y a trop de personnes qui/parce que ça c'est qu'est-ce que tu as dit
T: C'est pas c'que j't'ai dit Je regrette Mademoiselle

This example shows how confused students can be about the meaning of a word although they are able to associate a sensible interpretation with the context in which the word is learned (here the crowds and police that populate public ceremonies where protocol is used).

Consider now the kind of input that students receive concerning meaning. As already noted, there is a considerable emphasis on semantic information in the context of vocabulary instruction. Classes spent a lot of time during reading lessons going over the meaning of words. How this was done is interesting.

(14) No. 316

T: Qu'est-ce que c'est une casaque?
S1: Je sais pas Est-ce que c'est une/une ' apron '?
T: Non! Justement non! Ce n'est pas un tablier à ce moment-là
S2: Un chapeau
T: Non une casaque n'est pas un chapeau
S3: C'est un petit manteau noir
T: Bravo! Un veston, un manteau noir, n'est-ce pas? qui lui sert à représenter
son cocher

The students have clearly used various textual clues to figure out that the word in question refers to some kind of clothing (one might imagine that it is linked to the verb porter perhaps). Nonetheless, they do not know what the meaning of the expression is. S3 has figured out that the clothing represents the character's role when he's taking on the role of advocate (the excerpt is from Molière's L'Avare). The student therefore associates the item of clothing with the lawyer's black gown, which the teacher accepts. Now notice that this interpretation of the text is precisely what is needed to help everyone understand the play, but it is nonetheless true that the meaning of casaque is not "black coat". The dictionary definition which would be appropriate for the text is "coat with large sleeves". The colour is irrelevant but the shape is not. This definition is archaic. The word has a modern meaning which is "jockey silks". Here the colour is important and without further instruction regarding the status of the term, the student may misuse the term completely. One must be careful that in providing semantically-based instruction that instruction is not just instruction about context-bound interpretation. The instruction exemplified above was certainly appropriate to the situation and to the resource situations of the students but one does not want to find students at a loss with

(15)
Le jockey portait une casaque verte et blanche puisqu'il montait
pour la ferme de Lévèsque.

One can contrast the use of reading lessons to conduct vocabulary instruction with what occurs when the students read their own work. Here there was little intervention to provide the correct vocabulary. The following is illustrative:

(16) No. 336
S1: Elle savait que le dimanche était **la jour spéciale**
qu'elle ne devrait pas travailler
(bold text = un jour férié)
S1: Sa soeur ' Mary ' était **la seule personne qu'elle voulait dire**
qu'elle échappera bientôt
(bold text = sa seule confidante/la seule personne à qui elle
s'est confiée)
S1 Elle était **une tisserante**
(bold text = une tisserande/tisseuse)
S1: Madame Cook est tissée la plupart du journée à **faire la tissante**
(bold text = tisser)

This particular lesson offered many occasions to provide suitable vocabulary, most of them ignored. Occasionally, spontaneous discourse will provide similar opportunities.

(17) No. 316
S1: Peut-être il était un peu nerveux mais
il a **sauté beaucoup de mots**
T: Il ne les a pas sautés Il les a escamotés Qu'est-ce que ça
veut dire "escamoter"?
S2: Il ne les lise pas
T: À peu près cette idée
S3: Il a oublié
T: Exactement! Il oublie une petite partie ou une syllabe ou une
fin comme le mot "almanac"

Once again, the teacher is not satisfied to let the students use words which only approximate their intentions. Rather, the precise word is provided, along with a reasonable definition. Another example comes from class No. 116.

(18) No. 116
T: Est-ce quelqu'un pourrait me donner une définition
d' "expression idiomatique"
S1: des proverbes
T: Ce ne sont pas des proverbes
S2: une expression exagérée
S3: Tirer la vérité

T: Tirer la vérité exagérée la vérité c'est se servir des mots
pas dans leurs sens littéraire mais de se servir des mots pour
rendre une idée hein? Rendre une idée

Although one might sometimes question the accuracy of such off-the-cuff definitions, they do have a role to play in the classroom. They situate the expressions semantically and provide various paraphrases. Furthermore, one can add that this kind of instruction is in no way complicated; it does not demand extravagant preparation or costly materials. It simply requires that teachers understand fully that one of their major functions in teaching in or through a second language is to provide students with the words they need to say what they want to say. To give one final illustration of what can happen when one loses sight of this objective, consider (19):

(19) No. 356
S1: ...J'ai fait mon projet
T: Ton/ton projet sur quoi?
S1: ' starvation '
T: C'est quoi ça en français?
S2: (st ve n)
(...)
S3: malnutrition
T: C'est ça
S1: ' but that's malnutrition '
(...)
S3: Oui ' malnutrition ' et ' starvation ' sont co/ils sont euh un
peu la même
T: C'est pas la même chose toi?
S1: Non!
S3: Oui Malnutrition c'est...
T: Ah Name! Tu dis que c'est pas la même chose!
S: Oui
T: °C'est quoi les différences!°[30]
S1: °parce que° TU peux/C'est comme ' malnutrition ' et c'est/c'est
quand tu ne reçois pas ' glucose ' ' protein ' et ' carbohydrates '.
Tu peux manger comme ça et des ' chocolate bars well ' barres de
chocolat pour comme dix/dix jours et tu vas/tu vas avoir
' malnutrition ' mais ' starvation ' c'est quand °°tu ne°° manges pas
T: °°C'est°°/C'est quand tu meurs de faim °alors?°
S1: °oui°
S2: Oui
S1: ' malnutrition ' c'est quand tu ne reçois pas de
S2: des choses nécessaires
S1: Oui comme tu besoin ' glucose ' pour ton um..pour ton cerveau et
' protein ' pour euh ton ' liver or something '
T: Pour C'est quoi ' liver '
S2: °foie°
T: °Oui?°
S2: le foie
T: le foie
S1: et 'carbohydrates' je ne sais pas encore

This lesson is a prime example of what happens when students are allowed to talk about a subject which interests them. They communicate quite effectively but they cannot do so using the vocabulary of the L2. They do not know it. Thus there is a significant gap between their interlanguage lexicon and the L2 lexicon of standard French. Student 1 wants to make a sophisticated distinction between starvation and malnutrition and can do so only using the English words which he knows are not rendered in French by starvation and malnutrition. The discussion ends without anyone being told explicitly how to express these concepts in French. This consequence is unfortunate not just because of the lost opportunity to learn new words but also because the concepts are not rendered in exactly the same way in the two languages so there were additional vocabulary-related lessons to be learned. Thus English has two nouns for the concepts but French idiomatically renders starvation by verb phrases mourir de faim or crever de faim. There is a technical word inanition but this word is not commonly used. In contrast, malnutrition is rendered by a noun, namely sousalimentation. Learning this word could have led to a discussion of the uses of the verb alimenter and its derivatives pâtes alimentaires, aliments etc. As it is, the students can now draw the conceptual distinctions but they still cannot discuss them in French.

Formal instruction

Vocabulary instruction was sometimes observed within the grammar lesson, that is to say that instruction related to knowledge about words was taught then.[31] So, for example, in one class (no. 156), where the grammar lesson taught the difference between nouns and pronouns, students learned that certain verbs, e.g. neiger, pleuvoir, grêler, etc., have an arbitrary and obligatory kind of subject.[32] Thus the grammar lesson, whose focus was on transitivity and intransitivity, also conveyed syntactic (and in this case semantic) information about individual verbs.

Although much of this kind of teaching gets incorporated into planned lessons through the use of a grammar text, it is obviously true that teachers could and did exploit occasions that arose spontaneously to discuss the labelling of individual words. The example in (20) illustrates this quite clearly:

(20) No. 326
S1: Madame! Est-ce que "voir" est le même que "recevoir"?
T: Oui "voir" c'est le même que "recevoir" Regardez-moi un moment
avant de commencer l'écriture du/de la/du conjugai/de la
conjugaison Quel groupe est-ce verbe/le verbe "voir"
S2: La troisième groupe
T: C'est un verbe du troisième groupe Très bien Est-ce/Donc
c'est un verbe
S3: irrégulier
T: irrégulier Très bien Donc il change beaucoup

This kind of teaching would count as vocabulary instruction since the teacher conveys to the class that the two verbs in question belong to the same subclass of the class of verbs and that the subclasses are defined in terms of their formal characteristics (that is to say, by being "regular" vs. "irregular" in their conjugation form). The learning in this instance could be characterized as focussed but incidental; it is direct and "self"-initiated (since the student recognizes himself the resemblances between the two verbs and is asking for confirmation of the hypothesis that they are in fact the same type of verb).

Some limited emphasis on structural patterns was observed. There were examples of systematic instruction (invariably in the grammar lesson) focussing on the internal structure of words, and indirectly on rules of word formation. The following illustrate

(21) No. 116
T: "achat" est le nom.. le verbe?
S1: acheter
T: "étudiant"
S2: étudier
T: "serviteur"
S3: serviter
T: Ah! Là tu inventes des mots .. serv/si tu es un serviteur tu dois/
S4: servir
T: "lavage"
S5: laver
T: "grand"
S6: grandir
T: "gros"
S7: grossit/grossir

The purpose of this exercise is to teach the structural relationship between nouns or adjectives and the corresponding verbs. It thus establishes a certain paradigmatic relationship between the vocabulary items. Notice, however, that the exercise is rather limited. The selection of exemplars is done seemingly haphazardly since, e.g. lavage is the activity noun corresponding to the activity verb laver but étudiant is not an activity noun (the appropriate noun is étude). Again, the activity noun for the verb servir is service not serviteur. It is important that students know that some relations are not realized overtly by a suffix. One case was illustrated above, namely the relationship between participles (which belong to the class Verb) and derived adjectives (which occur in noun phrases, undergo agreement rules, etc.). In the example in (21), there is a contrast between one pair of words in a relation Verb:Noun — er:age but in another case the suffix -age corresponds to a 0 suffix in the relationship étudier:étude. The particular exercise fails to point this out and thus fails to draw the students' attention to the fact that these different formal pairs bear the same functional relationship. Furthermore, it is not made explicit that both étudiant and serviteur are secondary nominalizations, the first deriving in a regular way from a verbal participle, the second being only semantically linked to servir (its direct connection is to servitude via truncation of the suffix -ude and affixation of the agentive suffix -eur). The verb servir might have been linked to servant (like étudiant) and serveur (which derives by affixation of eur to the stem of servir). A more conscientious attempt to transmit structural information would not have led to the confusion which arises in example (21). Students expect a systematic relationship to be established by these kinds of exercises and text writers are not doing them a service by overlooking this fact. The error serviter is perfectly natural given the choice of previous items, since the exercise predisposes one to expect a simple morphological relationship between the items. Notice too that the teacher's response is particularly unilluminating for anyone who does not already know the correct response. This example thus shows a rather typical instance where teachers can and ought to provide explicit instruction about vocabulary. The semantic connection between servir, servant, serviteur, service, serveur, and servitude is an interesting one, one worth exploring. Furthermore, teachers should be made aware that even poorly organized exercises can be profitably used if certain kinds of errors are anticipated and the teacher takes the time to draw the appropriate conclusions with the students.

A final comment on the above example is worth making. How is the non-expert to identify suitable exercises to teach morphological information? Exercises should be kept very simple, with each question focussing on only one kind of relationship if the purpose of the exercise is precisely to draw the students' attention to form. There is little to be gained, it would appear, by mixing together several aspects of vocabulary organization. Thus, the question illustrated above should be split into at least three distinct questions (which could be taught at different times). The first question might try to establish the basis for comparisons between deverbal nouns and verbs or nouns and denominal verbs.[33] In this case the exemplars of one category would have the same meaning as the expression from which it is derived. A second question could then focus on secondary nominalizations where the secondary noun means something different from the verb it is related to.[34] A third question would elicit information on verbs and their corresponding adjectives. A final question might try to draw relations between nouns, verbs and adjectives all together.

The text in (22) provides an example of spontaneous instruction about morphological structure. The context consists of a series of activities combining the reading of a text (one of Aesop's fables) and an explanation of vocabulary items. The teacher, who has planned a lesson on the vocabulary of the reading, begins with a semantically oriented question about one of the items believed likely to be unknown.

(22) No. 326

T: un "botillon de paille" Qu'est-ce ça veut dire "un bottillon de paille"? D'où est-ce que ça vient ce mot-là "bottillon" Regardez-le un peu

S1: bouteille

S2: beau

S1: bouteille

T: "botte" "botte" Je prends une botte, vous savez, une botte de paille

S3: C'est un mor/c'est comme de la paille qui est mis dans un morceau

T: Oui Vous vous rappelez le film/pas le film!/le ballet qu'on a vu ensemble qui était "La Fille Malgardée"? Vous vous rappelez les paysans qui sont venus avec des bottes de paille sur la tête/euh sur l'épaule?

S4: Et il y a un/

T: Quelque chose qui les attache ensemble

S5: Une corde

T: Corde Bon! On arrête les/les commentaires (irrelevant text) mais alors/Pardon! le mot "bottillon" est plus petit que botte Comme vous vous rappelez vous avez fait un exercice "livre: livret", "porte: portière" euh

S1: maison maisonette

T: "maison: maisonette"

S2: ruisseau

T: "ruisseau: ruisselet"

S3: fille fillette

S4: garçon garçonnet

T: "nappe: napperon" Alors le mot "bottillon" c'est un petit/une petite botte de paille

This bit of text was discussed in a previous section devoted to an examination of how teachers use prior knowledge in the class. Here the focus is on the content of the instruction itself; namely morphological distinctions. To get at the meaning of the diminutive bottillon the teacher directs the students' attention first to the meaning of botte and then explicitly draws their attention to the formal relationship between the two words by recalling a previously done exercise on the formation of the diminutive. Only once the analogy is established does the teacher provide a definition of the word.

Another example follows hard on the heels of the first in the same lesson:

(23) No. 326
T: Qu'est-ce ça veut dire "une pincée"?
S1: un tout petit
T: Un tout petit peu! (irrelevant text) Faites-moi/Comment est-ce que vous dites une pincée de sel, oui? (student mimes action) Voilà!
Tu prends comme ça (gesture) avec les/les doigts
S2: un ' pinch ' (long exchange about significance of throwing salt over one's shoulder)
T: Ça (gesture) c'est une pincée de sel Comment est-ce qu'on appelle ça? (gesture)
S3: une poignée
T: Poignée Bravo! Une poignée, n'est-ce pas? Une poignée de eum de noisettes, une poignée de riz, n'est-ce pas? et une pincée de sel. Très peu, c'est-à-dire on/comme une pince

These exchanges are informative and spontaneous in that they reveal meaning in a systematic fashion and clearly indicate the function of the suffix which is the focus of attention. Furthermore, the teacher provides examples of uses of the words which are natural and spontaneous. They refer to and utilize the experiences of the students. The only minor improvement one might suggest to reinforce the structural comparisons would be to make explicit the relationship poing:poignée and pince:pincée as well as to mention other words illustrating the relationship (bouche:bouchée, cuillère:cuillerée, etc.). The idea would be to explicitly draw the learner's attention to the rule-governed nature of this relationship.

Finally, these exercises illustrate the functional identification of the morphemes since they are used to focus on the expression of a given meaning (rather than just on the form itself). The diminutive in example (22) is actually a morphological category rather than a semantic type[35] but it corresponds reasonably well to a concept, namely a small or cute thing. So exercises which teach that -ette, -et, -on all express this concept are not only teaching form they are teaching function or meaning.

The presentation of other types of grammatical information is virtually restricted to the official, planned grammar lesson. It is in this context that students learn to systematically group and identify words into syntactic classes and subclasses (e.g. first, second or third verb class, descriptive adjectives vs. quantifying expressions, stative vs. action verbs, etc.). There is little explicit teaching of this sort of information outside of the grammar lesson (although teachers may occasionally use the metalinguistic vocabulary that they have taught to their own students). One may conclude that a consequence of this is that there is little focussed learning on aspects of syntactic organization outside of the grammar lesson.

One fact which is striking about the presentation of words outside of the reading and grammar lessons is the extent to which items are introduced and manipulated independently of any syntactic context which might otherwise provide useful grammatical information. Consider the following:

(24) No. 136
S1: Qu'est-ce que c'est ' a lawyer '?
T: avocat

First, observe that the English word lawyer does not specify the sex of the referent but that in French one must do so if one knows it. Therefore, a fully informative response to the student's question would have been one which pointed out that male lawyers are called avocat while female lawyers are called avocate. A similar criticism can be made of the following exercise from the same class:

(25) No. 136
T: Encore quelqu'un peut me nommer des professionnels s'il te plaît
S1: médecin
T: médecin Un autre
S2: avocat
T: avocat etcetera

This kind of exchange occurs over and over, again where the students supply a one word response to their teacher's questions (with no indication of gender, for example) and the teacher naturally repeats the student's one word. These exchanges provide opportunities for expansions which would indicate to the student what the word's gender is. Text (26) provides an illustration in the same class:

(26) No. 136
S1: police ' arm force ' (English pronunciation) détective
T: Wo! Un policier Oui!
S1: un policier
T: ' arm forces ' An! An! (negation) "Un soldat" tout simplement Ensuite!
S1: euh détective
T: Un détective ça se rapporte à policier ça
S1: assassin?
T: un assassin?
S3: chauffer
T: un chauffeur
S2: cuisineur
T: un cuisinier Oui!

Clearly such expansions occur haphazardly over the course of the classroom observations. There are some empirical studies into the nature of error correction which suggest that it is done randomly (cf. Chaudron 1977, Fanselow 1977). There is some suggestion that haphazard correction is ineffective (Lightbown 1985). It is known from first language studies that explicit correction bearing on form has little effect on the child's production, that changes in behaviour are partially motivated by changes in the child's grammar. Second language studies suggest similar results, i.e. that unsystematic error correction is not likely to have an effect on the learner's developing L2 system even though the L2 learner is often older than the L1 enfant, and consequently is cognitively more mature. A rather amusing anecdote to confirm the tenaciousness of certain errors (whose exact cause still remains a mystery) is given in (27). This example comes again from class No. 326 who are still involved with the Aesop's fable. The students are reading out loud individually. There is ample graphic and linguistic evidence before the students concerning the gender of the words of the story. Nevertheless in reading, one child misreads a sentence as in (27a) continuing as in (27b):

(27) No. 326
a. S: Il alla directement s'adresser le sorcière
b. S: "Je suis venu t'apprendre une drôle nouvelle vieille sorcière"

The teacher ignores the first error. This is probably to be expected as there was no reason to assume that the student was confused about the interpretation of the text rather than merely making a speaking slip. A second student, however, repeats the same error in answering a question:

(28) No. 326
T: Comment est-ce qu'il a changé?
S1: Il a vu le bâton magique et il est peur
T: Il a peur maintenant
S2: Le sorcière peut faire quelquechose

Once again the teacher ignores the error which is repeated a third and a fourth time. Only then is there a correction.

(29) No. 326
a. S3: Qui dit "Cela ira très mal pour toi?" (reading question)
le sorcière
T: LA sorcière
b. S4: Les trois animaux faissent
T: faisaient
S4: faisient les fa/fanfarons devant le sorcière

It may be the case that despite overwhelming graphic, phonetic, morphological and semantic information to the contrary, the students believe that the referent is male, and that this error causes the erroneous use of le. This shows the importance of not ignoring errors. This is not to suggest the kind of disruptive and constant correction of minor points of grammar, but the systematic presentation of positive and negative evidence through the use of expansions and repetitions. This need not be confusing and teachers should also try to consider the manner in which they conduct their corrections. Suppose the teacher in the above examples had stopped the first student after the first error to ask if the witch was a woman or a man. By doing so the teacher could verify if the student (and indeed the whole class) understood the meaning of the word. One might have anticipated the following kind of dialogue

(30)
a. S1: C'est un homme?
T: Alors ce serait "le sorcier" et non pas "la sorcière" tout comme c'est "un écolier" mais "une écolière", "un fermier" mais "une fermière"
b. S1: C'est une femme
T: Alors on dit /la/ sorcière tout comme on dit /la/ fermière /la/ bergère, etc.

It is true that gender problems can be sorted out more readily with words where gender corresponds to a sex-based distinction in the reference (although there is by no means a simple connection in French between a referent's sex and a noun's gender). Most words, however, do not function this way. Nevertheless, this still does not preclude systematic but skilled interventions to draw the students' attention to what they are not picking up from mere isolated exposure to words.

Teaching sociolinguistic information

It was noted in previous sections that vocabulary comes with its own identity. Words have a history, they move in certain social circles and they have if not a nationality at least a territory. Students must be made aware of restrictions on the use of words. This is particularly important when they are reading texts from another era. One of the classrooms was reading Molière's L'Avare (see example 14). It was noted there that the meaning of the word in the text was not the same as its modern-day meaning. This was not pointed out, which is unfortunate. Such lapses leave open the possibility of misuse in the future. More importantly, students lose the opportunity to see how words can change meaning (a topic which in itself would be an interesting part of any "grammar lesson"). Turning specifically to sociolinguistic information, there was no evidence from the observations that teachers were consciously attempting to teach styles or registers which might be associated with one particular social group. Spontaneous teaching does occur, as in (31) below, but observe that the differences between the two uses are not explained

(31) No. 136
S1: ' steak ' Comment on dit ' steak '?
T: Il y a des gens qui disent "steak" Il y a des gens qui disent "bifteck"
Enfin j'accepte "steak"

While the information that the teacher provides here is accurate, it is so vague as to be useless to the student one wants to know when steak would be appropriate and when bifteck would be. One is certainly more current than the other in Canada (namely steak) but only bifteck is found in commonly used dictionairies. Example (19) also illustrated the same type of problem, i.e. of the need to provide information about the proper context of use for a given word. Thus, it was pointed out that French has a noun form to mean 'starvation'. The English student may be predisposed to expect a noun to express the concept because English has one. However, French only has a noun in technical language. One could say to students "Doctors, nurses and biochemists say inanition but in ordinary conversation one says mourir de faim."

On a separate but related issue, no teaching on specifically discourse-related functions of words was observed. Teaching was obviously focussed on coherence (as noted repeatedly above) but particular cohesive functions of individual items were not discussed.

2:8 Conclusions

The analysis of vocabulary teaching activities in Grade 6 immersion classes indicates that it occupies a rather narrow place in the overall teaching plan. This conclusion is warranted by the strong link between planned teaching and reading, with the necessary limitations on choice of words that follows from a focus on the written medium (formal language, context independent language etc.). Teachers and text writers appear to be underestimating the central place of word knowledge in determining both grammatical knowledge and communicative proficiency and hence are not giving enough prominence to the place of vocabulary instruction in relation to other kinds of activities.

The conclusion is also justified by the limited linguistic focus of instruction. Teachers teach meaning, more often interpretation. Thus they focus on the precise interpretation of words in a specific context without also providing exercises or instruction about the use of the words in other contexts. This is not sufficient if one objective of instruction is to help students learn to go beyond immediate contexts to more general situations. Teachers and texts could also focus more frequently on phonological, morphological and syntactic aspects of vocabulary knowledge. The "grammar lesson" should be seen to include a vocabulary component. Indeed, one might consider abandoning organized, overt grammar-focussed activities except to provide metalinguistic vocabulary for the purposes of explicit linguistic analysis (e.g. sujet, nom, verbe, prédicat, attribut, etc.). The contents of grammatical analysis can be taught in any number of ways and it is possible that more linguistically relevant information might be conveyed if that content became the focus of all types of lessons. Finally, learners must be taught sociolinguistic and discoursal aspects of word use. These aspects of proficiency are also too infrequently the focus of planned instruction in grade 6 French immersion. The analysis suggests that teachers are doing the right things but that they are not doing them with regularity and systematicity. Improvements could arise from simply being conscious of the need to help learners build a better dictionary and making more of an effort to build classroom activities around this need.

3. VOUS/TU INPUT

The French pronouns tu and vous carry both grammatical and sociolinguistic information. A number distinction may be signalled by the use of singular tu versus plural vous. A sociolinguistic distinction may be manifested in familiar tu versus formal vous, a marker of respect or politeness.[36]

In the second year report of this project, it was reported that on sociolinguistic measures of the use of vous and tu, grade 6 immersion students tended to overuse tu in situations calling for the use of vous (e.g. making a request to an adult). That is to say, in formal contexts, vous was underused by immersion students relative to native speakers of the same age (pp. 40-45). Swain (1985) suggested that the underuse of vous as a politeness marker in formal contexts could be linked directly to the input the students received. The input, she hypothesized, would reveal the singular/plural distinction between tu and vous as teachers would use the former to address individual students and the latter to address groups of students or the class as a whole. It was thought, however, that input relevant to the sociolinguistic distinction would likely be infrequent since there seemed to be few opportunities for the teacher in the classroom setting to naturally make use of vous as a marker of deference.

In a separate study of grade 10 early and late immersion students, Swain and Lapkin (in press, see Appendix C, pp. 293-316) also examined the use of vous and tu. They used the DBP sociolinguistic oral measure as well as other communicative measures. As with the grade 6 early immersion students of the DBP study, they found that the grade 10 early immersion students tended to overuse tu in formal situations relative to native speakers. Late immersion students, however, behaved quite similarly to native speakers in formal situations but tended to overuse vous in informal situations. In order to explain their results, Swain and Lapkin speculated that early immersion students were 'doing as their teachers do, not as their teachers tell them to', while the late immersion students were 'doing as their teachers tell them to, not as their teachers do'. In other words, it was assumed that both groups of students were receiving 'sociolinguistic instruction' about the appropriate use of tu and vous, but that early immersion students were continuing to base their production on early acquired rules.

The transcripts from the recordings made in the early immersion classes provide the opportunity to verify these hypotheses as they pertain to the nature of the input received by these students. It was decided to examine the data from the grade 6 classes only, as that was the level where production data were also available.

All uses of tu and vous by the ten grade 6 teachers were counted, as were the uses of the same pronouns in the public talk of the students. Not counted were the uses of vous and tu by students in performing formulaic exercises. For example, in one class, students were asked to transform sentences from the 2nd person singular to the second person plural: such uses were not included in the count on the grounds that the students' responses were predetermined rather than spontaneously uttered. There was also considerable evidence that such tasks were undertaken without paying attention to the sentence's meaning, as indicated by the following example:

(1) No. 156
T "et que tu nous aideras à préparer les décors de notre fête"
S et que et que vous nous aiderez de préparer
T oie yoie
S1 et que nous vous aideriez?
S2 nous aideriez nous ai/
T ah/ oie yoie
S1 et que vous nous aiderez

Later, it became apparent that the teacher had been reacting to the students' lack of liaison between nous and aiderez, but the students in the meantime randomly reordered pronouns and tried out different verb endings.

The pronouns were classified according to the functions they served: singular, plural or generic; formal or informal. An example of each is given below.

(2) singular and informal tu
T Tu peux l'avaler ou bien le jeter dans la poubelle.

(3) singular vous
T (Name) un s'il vous plaît et vous l'épelez

(4) plural tu
T la feuille que je te passe, tu la prends et..

(5) plural vous
T Vous allez me dire s'il y a des arguments

(6) generic tu
T Si tu es un serviteur, tu dois..?

(7) generic vous
T Et vous avez plusieurs verbes qui ne changent pas

(8) formal vous
T Madame, voulez-vous un livre?

In isolation, the classification of these pronouns may seem arbitrary. However, in context, it is usually quite clear. The initial classification was done from the transcripts by a native speaker of French. The pronouns from three of the transcripts were independently classified by a fluent speaker of French. Few disagreements occurred. Those that did occur were reconciled by listening to the original tapes.

As indicated in Table 4 (p. 253), tu and vous were used about equally often: an average of 89.3 times (52.7% of the time) and 80.2 times (47.3%) respectively. Tu was generally used by teachers when addressing individual children. In eight of the ten classes, however, tu was occasionally used by the teachers to address the class as a whole ($\overline{X} = 6.3$). Of the uses of tu, approximately 7% were used in plural contexts. A reasonable inference based on this input is that tu is not restricted to the singular.

Table 4 also indicates that vous is typically used to address the class as a whole. However, occasionally vous was used by the teachers to address individual students ($\overline{X} = 1.6$). Thus, it appears that the input the immersion students hear indicates that vous as well as tu can function as both a singular and a plural pronoun.

This is not the message conveyed by their instruction, however. Their instruction appears to focus on the number distinction. For example, in several of the classes during the observational period, students carried out various exercises which focussed their attention on the number distinction between tu and vous.

(9) No. 156
T Quelle est la première personne du singulier, pronom personnel?
S Je
T Je. OK. Du pluriel (name)?...première personne
S Nous
T Nous voilà. Deuxième personne continue (name)
S Celle ou celles à qui l'on parle
T Celle ou celles à qui l'on parle. (name). Deuxième personne du singulier
S Tu
T Deuxième personne du pluriel (name)?
S Vous

Other examples involved changing entire passages from the singular to the plural: "Mettons au pluriel le passage suivant"; or from the plural to the singular: "Modifions les phrases suivantes en employant la deuxième personne du singulier".

Table 5, p. 254, indicates that there are very few occurrences of vous where it is used as a politeness/deference marker by teachers. There is less than one such occurrence ($\bar{X}$ = 0.2) per class. Examples typically involve an exchange between a teacher and an adult visitor to the class.

These results confirm the original hypothesis that early immersion students tend to underuse vous as a politeness marker, not because the form vous is infrequent in their environment, but rather because its function as a sociolinguistic marker is infrequent in the classroom context.

But the results also suggest that more is involved in forming the patterns of student usage of vous and tu than the relative infrequency of the use of vous as a sociolinguistic marker. The input to the students provides them with evidence that both tu and vous can function as a singular or a plural pronoun, as well as a generic pronoun: two second person pronoun forms might well be considered redundant. This, coupled with the fact that in their native language there is only one form for the second person pronoun, you, could foster the use of only one form in French.

Why would that form be tu rather than vous? There are at least two possible reasons. First, by using tu, one can avoid the complexity of the verb forms associated with vous. Second, it might be that opportunities for the obligatory use of vous by the students to signal plurality are rare in the classroom setting.

Evidence from the students' public and spontaneous use of vous and tu noted in the transcripts strongly supports the second reason. (The transcripts cannot directly address the first reason.) Table 6, p. 255, shows that tu is the overwhelming choice of students in their use of second person pronouns: overall tu is used 96% of the time ($\bar{X}$ = 21.6) and vous is used only 4% of the time ($\bar{X}$ = 0.9). Table 6 also indicates that there is very little use of either tu or vous in plural contexts — a total of 1.3% ($\bar{X}$ = 0.3). When the occasion does arise requiring the use of a second person plural pronoun, tu is twice as likely to be used as vous ($\bar{X}$ = 0.2 and 0.1 respectively). What this means is that occasions for the obligatory use of a plural, second person pronoun are rare in the normal flow of classroom interaction. Furthermore, when given the opportunity, students are more likely to use tu.

Table 7, p. 256, provides corroborative evidence that the grade 6 immersion students are overusing tu in contexts requiring the use of vous as a sign of politeness/deference.[37] Tu is regularly used to address the teacher. The following represent several examples.

(10)
S Madame tu n'as pas venu.

(11)
S Monsieur, où est-ce que tu peux avoir les/les choses qu'on fait pour le
T Pour l'école d'été? euh au bureau en arrière.

(12)
S Je/tu veux que je juste lis ça?
T Non. J'veux que tu mettes le passage au pluriel

Thus, as suggested, the early immersion students are 'doing as their teachers do' to the extent that they are provided opportunities to do so. But the classroom environment seems impoverished in two ways. First, opportunities for students to observe the sociolinguistically motivated uses of vous and tu in regular classroom discourse appear limited. Secondly, opportunities for students to use the grammatically motivated use of vous also appear to be infrequent in regular classroom discourse. The observations argue, once again, for the introduction of instructional materials and/or activities which fill these gaps.

A hypothesis to be pursued is that, at this age and with the knowledge the students already possess, provision of relevant grammatical and sociolinguistic rules in context, along with adequate opportunities for obligatory use, would benefit learning.

4. STUDENT TALK IN TEACHER-FRONTED ACTIVITIES

In attempting to explain the French proficiency results of grade 6 immersion students obtained in the second year of the DBP project, Swain (1985) argued that the learning problem could not be accounted for by the lack of comprehensible input, but rather might be accounted for, in part, by the lack of demands for 'comprehensible output'. 'Comprehensible output' was defined as producing language in situations where the learner is pushed towards the delivery of a message that is not only conveyed, but that is conveyed precisely, coherently and appropriately.

It is important to note that there are two aspects to the concept of comprehensible output. First, there is the notion that language needs to be produced to be moved towards native-speaker proficiency. Secondly, there is the notion that feedback needs to be provided to learners in order for them to make progress in developing their linguistic systems. The second of these raises the question of how teachers respond to the errors made by immersion students, and will be treated in the next section on 'error treatment'.

The first point, that language needs to be produced to be fully acquired, needs to be explored further. There are at least two roles in second language acquisition that might be attributed to output. One, as Schachter (1984) has suggested, is the opportunity it provides to test out hypotheses — to try out means of expression and see if they work.

A second function is that using the language, as opposed to simply comprehending the language, may force the learner to move from semantic processing to morphosyntactic processing. As Krashen (1982) has suggested: "In many cases, we do not utilize syntax in understanding — we often get the message with a combination of vocabulary, or lexical information plus extra-linguistic information" (p.66). Thus, it is possible to comprehend input — to get the message — without a syntactic or morphological analysis of that input. The claim, then, is that producing the target language may be the trigger that forces the learner to pay attention to the means of expression needed in order to successfully convey his or her own intended meaning.

An even stronger claim can be made on the basis of the proposals of Clark and Clark (1977) and van Dijk and Kintsch (1983) for the notion of a fuzzy, open, nondeterministic syntactic parsing strategy that can be used for comprehending discourse but would be inadequate in producing it. It may, therefore, not be just that only semantic processing is required for comprehension but that in addition any syntactic processing involved in comprehension might be quite different from the closed logical system of rules required to produce a grammatical utterance.

An assumption behind Swain's (1985) proposal that immersion students' grammatical performance would be enhanced with additional opportunities to produce French was that few such opportunities were provided in class. The grade 3 and 6 immersion observational data provide a means of examining this assumption.

In this section, opportunities students have to speak French are examined. The opportunities students have to write French are not considered directly, but only as they provide the basis for students' reading aloud of their own work. It is, however, worth noting that much of the writing the students do is of the 'exercise' form: writing out fairly narrowly prescribed responses to teachers' questions based on the reading of a text (e.g. history) — the answers can frequently be copied directly from the text; writing out sentences that contain a grammatical point being taught; or filling in blanks. Few examples were observed of creative writing, and no examples were observed of journal writing.

4:1 The Data Base

Three sets of data are used in considering student talk. The first set consists of all the French spoken by grade 3 students in each of the nine immersion classes that was audible enough for transcription in a 90-minute period of time (both sides of one tape). As most private conversations between teacher and student, or between student and student, were inaudible, the data largely represent student talk that was directed to the teacher in teacher-fronted activities. The second set is similar except that it represents the French spoken by grade 6 students in each of the ten immersion classes. For the grade 3 students, the 90 minutes account for less than half of their daily instructional time in French; whereas for the grade 6 students, the 90 minutes account for most of their daily instructional time in French. The third data set consists of all the transcribable English spoken by the grade 6 students in the English portion of their school day. The amount of time varied from 53 minutes in one class (No. 346) to 112 minutes in another class (No. 316). The average number of minutes per class, however, was 88.

Each student turn was categorized, with several exceptions. If a turn was interrupted by the teacher or another student, but the student continued on as if there had not been an interruption, the entire utterance was counted as one turn. For example:

(1)
S Il faisait du vélo et alors il a tombé parce que
T il est tombé
S parce que la roue...

Also counted as one turn were instances where the student talk is obviously part of the same turn, as in (2) below.

(2)
S Madame
T Oui
S quel projet est-ce que ça sera?

If the student repeated his or her words, the repetition was not counted. Thus (3) would be counted as only one turn. If there was a lexical or syntactic change in the repetition, the repetition would count as a second turn.

(3)
S Madame est-ce qu'on écrit?
T Pardon?
S Est-ce qu'on écrit?

Not counted were non-language comments such as "sh-sh-sh", "aw", and "uh-oh". ("Yeah", "OK" and "hey" were counted.) Turns incorporating untranscribable words or sections were coded in the appropriate categories only if the structure remained clear. Otherwise they were excluded from the counts.

4:2 Classification of Student Turns

Student turns were classified in two ways: length and source of talk.

First, student turns were categorized according to their length as either minimal (M), phrase (P), clause (C), or sustained (S). The last two categories were, for some analyses, added together and labelled as extended (E). These categorizations were made and counted for both the English and French transcripts.

Minimal length refers to turns of one or two words in length, or a list of words. Spelling a word was also considered to be a minimal turn. Other examples included "madame", "oui", "non", "yeah" and "OK".

A turn was categorized as P when it consisted of an adverbial phrase ("aux Antilles", "par le professeur"), a nominal phrase ("le garçon", "un très bon acteur") or verb phrase ("est montrée", "pas ménagère", "regarde la télévision"). Also included as phrases were such mathematical routines as 'a number times a number' ("vingt fois six"). A turn that consisted of M + P was classified as a P, for example, "non, des monstres" or "OK, à vingt-deux heures".

A turn was categorized as C when it consisted of a simple clause such as "parce que il n'y a pas d'action dans le programme" or "je pense aujourd'hui". A turn that consisted of M + C was also classified as a C. Examples include "Madame, je suis prêt" and "oh, je sais".

Sustained talk was considered to be any turn that was more than C. Examples include "est-ce que tu sais quand il vont avoir les résultats?" and "comme quand/quand ils commencent de... vivre aux Antilles, ils doivent avoir des 'sclaves et ça change parce qu'ils n'avaient pas d'esclaves avant".

The second way in which student turns were classified was according to the source of the talk. Student talk can be student-initiated, either by themselves or in response to another student's utterance. Or, student talk can be teacher-initiated. The teacher can initiate student talk by different means: asking a question, giving an order, naming a student. Within teacher-initiated talk, the student can be reading aloud from his/her own written production or spelling aloud from his/her own written production. These two categories were considered as pre-planned (by the student) teacher-initiated student talk. It was also possible that the talk was not pre-planned (by the student). Unplanned teacher-initiated student talk was classified into seven categories: repeating (prompted or not), spelling, completing someone else's (usually the teacher's) utterance, selecting from limited linguistic choice (yes/no; answers that are semantically controlled — verb

forms, parsing, mathematical operations), finding own words, reading aloud, and reciting from memory. These categorizations were made only for the French transcripts. Examples from each of these categories are given below.

Student-initiated talk: self

(4) S Qu'est-ce que c'est "vaniteux"?

(5) S Madame est-ce qu'on écrit?
T Pardon?
S Est-ce qu'on écrit?

(6) T Mais j'n'ai pas fini de corriger encore.
S Est-ce que les personnes a eu mal ou bon?
T Je n'ai pas encore...je...je suis encore en train de corriger des sections. Je n' peux pas te répondre.
.
.
.
S Tu mets combien de temps?
T Combien de temps? Ça prend beaucoup de temps. Ça prend quinze minutes par projet.

Student-initiated talk: other student

(7) S1 J'ai pris 'wood carving'.
S2 Moi, je n'ai pas pris ça.

(8) S1: Est-ce que 'Bishop' était sur 'Innes Road'?
S2 Je n'va pas là maintenant!

Teacher-initiated student talk: pre-planned, reading aloud from own writing

(9) T Vas-y.
S Quand j'étais petit un chose très drôle a passé. C'est l'hiver, il faisait très froid. J'étais en colère à quelque chose ma mère a fait. En ce temps-là, j'avais mis un paire de shorts et un T-shirt/T-shirt?
T Oui.
S J'ai venu en bas et j'ai dit à ma mère "Je n/je pars et je ne reviens plus". J'ai frappé la porte derrière moi très fort. C'était très/c'était très froid dehors alors j'ai revenu dans la maison et j'ai dit à ma mère à un voix méchant "Je vais te donner une autre chance".

Teacher-initiated student talk: pre-planned, spelling from own writing

(10) T (name) un s'il vous plaît et vous l'épelez.
S donne. D-O-N-N-E.

Teacher-initiated student talk: unplanned, repeating

(11) S Est-ce que je vas euh aller à onze heures à le bureau du...du eum "nurse"?

T De la garde.
S De la garde.

(12) S Et voilà, ils ont battu la vache.
T Ils ont abattu.
S Ils ont abattu la vache.

Teacher-initiated student talk: unplanned, spelling

(13) S Il va manger.
T Épèles 'manger'.
S M-A-N-G-E-R.

Teacher-initiated student talk: unplanned, completing

(14) T Pousser comme de la mauvaise...?
S herbe.

(15) T Quelle est la question qu'on doit garder en tête?
S Le changement.
T Tu l'as. Le changement...?
S1 du monde.
S2 dans les années.
S3 de leur vie.

Teacher-initiated student talk: unplanned, selecting from limited choice - yes/no

(16) T Est-ce que ça vous arrive parfois?
S Oui.

Teacher-initiated student talk: unplanned, selecting from limited linguistic choice - semantic

(17) T Quel est cet objet?
S Une assiette.

(18) T Est-ce que quelqu'un pourrait me donner une définition d'expression idiomatique?
S Une expression exagérée.

(19) T Donne-moi la définition.
S Un adjectif qualificatif est un adjectif qui décrit une personne, les choses.

Teacher-initiated student talk: unplanned, selecting from limited linguistic choice - syntactic

(20) T Première personne du pluriel?
S Nous

(21) T Qu'est-ce que je fais maintenant (name)?
S Six plus huit plus zéro est quatorze.

(22) T Tu poses?
S Tu poses le quatre et retiens le un au-dessus de...non...umh

Teacher-initiated student talk: unplanned, finding own words

(23) T Qui pourrait décrire la situation?
S Elle a réveillé en bas du lit.

(24) T Tu peux venir nous parler de ta fin de semaine
S Samedi matin j'ai regardé la télévision et samedi après-midi j'ai allé m'acheter...

Teacher-initiated student talk: unplanned, reading aloud

(25) T Lis la phrase.
S J'ai bien mangé ce matin.

(26) T Qui n'a pas lu dans cette classe? (name), vas-y.
S "D'accord" dit le lion "mais pas une vache de mon troupeau". "Pas un des miens" dit l'éléphant. "Et du mien non plus" ajouta vite le léopard.

Teacher-initiated student talk: unplanned, reciting from memory

(27) S Le boulanger. Qu'il est drôle le boulanger avec ses cheveux couleur de farine sur ses mains sur ses bras sur sa poitrine. On dirait qu'il vient de neiger. Sans se lasser d'un geste prompte tandis que/tandis qu'au village chacun sommeille...il moule des pains en creux des corbeilles pareilles à des chats accroupis en rond. Puis dans le four au coeur vermeille...il les plonge au bout d'une longue pelle et bientôt miches en ribambelles sortiront couleur de soleil.

The French transcripts were classified by a native speaker of French and the English transcripts were classified by a native speaker of English. Both individuals have training in linguistics. One transcript from each language was coded by a third person, a native speaker of English. The final counts in each category rarely differed by more than a frequency of two. This was considered satisfactory as a reliability check; therefore no further checks were made.

4.3 Results

The results are displayed in Tables 8 through 12 (pp. 257 - 261). Tables 8 through 11 are concerned with student talk in 90 minutes of the French portion of the instructional day. Table 8 shows the frequencies for each category of student talk for each of the nine grade 3 immersion classes. It also shows the average percent for each category. Table 10 contains precisely the same information, but for the ten grade 6 immersion classes. Tables 9 and 11 break down the category "selects from limited choice" into its three subcategories -- yes/no; responses semantically controlled; responses syntactically controlled -- and show frequencies for each class, average frequencies and average percentages. Table 9 is concerned with the grade 3 classes; Table 11 with the grade 6 classes. Table 12 includes data from both the English and

French portions of the day. It summarizes data presented in Tables 8 to 11 concerned with the length of student turns. It also expands the category 'extended' into 'clause' and 'sustained'.

Tables 8 (p. 257) and 10 (p. 259) reveal a considerable similarity between grades 3 and 6 in terms of the sources of student talk. Two differences worthy of note are that the grade 3 students appear to spend more time repeating the teachers' utterances than do grade 6 students (9.9% vs. 5.1%), while the grade 6 students spend more time finding their own words than do grade 3 students (22.9% vs. 17.9%). Tables 8 (grade 3) and 10 (grade 6) indicate that the most frequent source of student talk is 'selecting from limited choice' (36.5% and 37.8% respectively), which appears to encourage minimal responses (23.6% and 25.0% respectively). The second most frequent source of student talk is "finding own words" which appears to encourage extended responses (13.7% and 18.3% in grades 3 and 6 respectively). Self-initiated responses also appear to encourage extended responses (16.0% and 11.9% respectively). There is slightly more student-initiated talk in grade 3 (22.4%) than in grade 6 (17.6%).

As indicated by Tables 8 and 10, 'selecting from limited choice' is the most frequent source of student talk. When this category is broken down into its subcategories, the most frequent for grade 3 are responses that are syntactically controlled (36.5% vs. 31.8%) whereas for grade 6, responses that are semantically controlled are more frequent (46.7% vs. 22.0%) (see Tables 9 and 11). In grade 6, semantically controlled responses also produce a higher mean percent of extended responses (16.4% semantically controlled vs. 5.5% syntactically controlled). In grade 3, however, semantically and syntactically controlled responses produce roughly equivalent mean percents of extended responses (14.4% and 13.7% respectively).

Table 12 (p. 261) shows that the average number of student turns per class in grade 3 is 223.2 and in grade 6 is 226.7. Given that these figures represent 90 minutes of instructional time in French, a rough approximation of the frequency of student talk is two and a half times per minute. In the English portion of the day, student talk is approximately one and a half times as frequent, amounting to about four student turns per minute.

Table 12 indicates that when extended talk in French is broken down into talk of clause length versus talk that is more than a clause in length (sustained talk) and when sustained talk that consists of reading aloud is subtracted, less than 15% of all student turns are greater than a clause in length. The percent is slightly higher in the English portion of the day (16.7%). This percent, however, reflects one and a half times the number of student turns found in the French instructional time.

4.4 Discussion

This section on student talk began with the hypothesis that immersion students do not have opportunities to engage in 'comprehensible output'. In this section, the concern is solely with the 'output' part of the two-part notion of 'comprehensible output'. As will be seen in the next section, teachers do very little to make students' output more comprehensible, and what is done is thoroughly unsystematic.

With respect to the question of student 'output', the data show that students have opportunities to speak in French, but it is less than two-thirds of that provided in classes where the instruction is in their native language and where they also have the support of the out-of-class environment for language development. The 'output' hypothesis would suggest that second language learners need at least as many opportunities to use the target language as native speakers. Yet immersion students are not only getting fewer opportunities out-of-class, but in class as well. .

One way of increasing the opportunities for students to talk is to involve them in group work. For teachers who claim that this is not possible to do with immersion students because "they will talk in English", there are solutions. A well-known second language pedagogical solution is to make sure that the students have a task that **requires** the outcome to be a text in French -- spoken or written. Examples include the preparation of books to be published for classroom consumption, the preparation of a student newspaper, the preparation of a radio show, etc. Students preparing for a known audience will willingly spend hours working through self-generated or group-generated texts in order that they meet firm standards of perfection. This represents 'comprehensible output' at its best.

Almost no group work was observed in the immersion classes. In fact, in one of the few instances where students were grouped together, the teacher specifically prohibited the students from talking to each other:

(28) T Vous allez grouper ceux qui ont pas de livre avec ceux qui ont le livre.

.
.
.

T Euh, j' me demande si on a besoin de parler pour ça. Hein? (Name) et (name)? J' me demande si on a besoin de parler pour ça.

The results show a considerable similarity between grades 3 and 6 with respect to the sources of student talk. They also indicate the sources that are most likely to lead to extended talk, and the relative paucity of sustained student talk. The results suggest that by providing more opportunities for student-initiated talk and by teachers asking more open-ended questions, the amount of sustained talk would increase. Given that immersion education is an education that aspires to the two instructional goals of second language learning and content learning -- both of which profit by students talking -- then the results of this study suggest the need for considerable change in the French portion of immersion classes.

5. ERROR TREATMENT

As we saw in the last section, the notion of 'comprehensible output' has two components: that learners need opportunities to produce the second language; and that they be provided with feedback that will 'push' them towards a more coherent and accurate production of their second language. This section is concerned with the latter aspect: specifically, with French immersion teachers' verbal reaction to errors in their students' spoken French.

In current theories of second language acquisition, errors are regarded positively as signs of the developing interlanguage of the learner. Errors reflect the rules which underlie the use of the language by the learner. This does not imply that errors should not be corrected, but rather that error correction should be aimed at helping learners develop and test hypotheses about target language regularities, and learn exceptions to these rules. It is fair to say that how this is to be done is a matter of considerable interest, and one for which few answers exist.

Within the context of communicative language teaching, a dilemma exists: if teachers correct student errors, then the flow of communication is at least temporarily stopped, if not halted completely. If teachers do not correct errors during the flow of communication, then opportunities to make crucial links between form and function are reduced. Chaudron (1986:47) comments: "The multiple functions of feedback, as reinforcement, information, and motivation, and the pressure on teachers to be accepting of learners' errors lead, however, to the paradoxical circumstance that teachers must either interrupt communication for the sake of formal TL correction, or let errors pass 'untreated' in order to further the communicative goals of classroom interaction."

However, recent studies (e.g. Long 1980, Pica and Doughty 1985, Varonis and Gass 1985) have shown that in conversations between native speakers and non-native speakers, or between two non-native speakers, error correction that furthers conversation can occur through the use of such devices as clarification requests, confirmation checks and repetitions. What has not yet been demonstrated is the extent to which second language acquisition occurs as a result, although there is some indication that certain types of responses are more likely to elicit a correct form from learners than are others. For example, Chaudron (1977a), in a study of teacher corrections in a late immersion program, compared the effects of different types of repetitions: simple repetitions versus repetitions with emphasis, reduction or expansion of the learners' errors. He found that when the teachers repeated the students' errors with emphasis (either questioning tone or stress) or with reduction (i.e. repeats correctly the incorrect portion of the student utterance), there was an increase in correct responses given by the students. Chaudron concludes that identification of the error and/or unambiguous, explicit provision of a correct model can result in immediate, if not long-term, correction by learners of their errors.

In this section, the type of error treatment is not investigated quantitatively. Nor is it examined which type of error treatment elicited more correct responses from students. Both of these questions are of interest, and when the opportunity arises in the future to carry out further in-depth quantitative analyses of the data, it is our intention to consider these issues. However, the examples provided will illustrate immersion teachers' use of a variety of error correction strategies, including those of repetition, clarification requests and confirmation checks.

In this section, the focus is two-fold: first, what errors do teachers correct? and second, how systematic are their corrections?

The first question -- what errors do teachers correct -- has been examined in several studies (Chaudron 1977b, Courchêne 1980, Fanselow 1977, Lucas 1975, and Salica 1981). In summarizing their results, Chaudron (1986) notes that the median percentage of phonological errors corrected was 54% ($\bar{X}$ = 48%) and of grammatical errors corrected was 49% ($\bar{X}$ = 51%). These figures are of interest as it will be seen that they differ substantially from those in immersion classes. The specific questions addressed in this section are whether certain categories of errors are proportionately more subject to correction than others, and what the overall rate of correction is. This provides a quantitative context in which to examine qualitatively the systematicity of error correction.

5:1 The Data-Base

The data used for the tabulation of student errors and teacher responses were the complete transcripts of the ten grade 6 immersion classes. The average length of time recorded was 106 minutes, which accounts for the entire portion of the school day devoted to instruction in French.

Errors in surface level grammar and pronunciation were noted and classified. Content errors were not tabulated.[38] Also, lexical errors, which were considered in section 2 above, were not counted. For each error, it was noted if the teacher corrected the error. An error was counted as being corrected whether the correction was explicit (see (1) and (2) below) or implicit (see (3) and (4) below).

(1) S Il avait vu, nous avons vu
T Non! pas avons, avions vu. A-V-I-O-N-S vu.

(2) S Un lampe.
T Une lampe.

(3) S Parce que les aiguilleurs du ciel peut pas.
T Bien, les aiguilleurs du ciel ne pourront pas.

(4) S Le date aujourd'hui?
T Euh la date, c'est le huit.

5:2 Error Classification

The errors were classified as follows:

Gender assignment. Incorrect gender is assigned to a noun, pronoun or adjective.[39]

(5) S J'ai un question

(6) S (la princesse) Il devrait viver avec lui.

(7) S Sa père.

(8) S Tous les différents cultures.

Article assignment. An article occurs where there should not be one, is missing, or is incorrect in discourse. Note that incorrect number assignment was classified as an article error, whereas incorrect gender assignment was classified as a gender error.[40]

(9) S Il y a des certaines choses.
T Oui, certaines choses.

(10) T Qu'est-ce qui sont disposés?

S Devoirs.
T Les devoirs.

(11) S Et puis il voit une petite ferme. Il entre dans une ferme.

Contraction. A contraction is made where it should not have been, or it is not made where it should have been. Note that if there was an instance where a preposition and article were contracted erroneously, it was arbitrarily assumed to be a case of erroneous contraction rather than erroneous gender assignment. Thus, for example, (12) and (13) were classified as erroneous contractions rather than as gender errors. Examples (14) to (16) illustrate other contraction errors.

(12) S Le projet était du discrimination.

(13) S C'est pas du gomme.

(14) S Et c'était à le cottage de mon père.

(15) S Qu'il jette les petites choses à les fenêtre.

(16) S Le auteur
T Le nom de l'auteur.

Pronouns. This category includes missing pronouns and the incorrect choice of pronouns. The latter includes pronouns with incorrect number assignment, but not with incorrect gender assignment which were categorized as gender errors.

(17) S Je sais le poème mais je ne veux pas faire.

(18) S Le lièvre je ne voulais pas.
T Il voulait pas.

(19) S Elle voulait lui marier.

Prepositions. This category includes the addition of a preposition where there should not be one and the incorrect choice of a preposition.

(20) S C'est sur les mercredis.

(21) S Quand ils commencent de vivre aux Antilles.

(22) S S'est marié à Mary.
T Avec elle.

Verbs. Verb errors include wrong choice of auxiliary, missing verbs, wrong tense, wrong number or wrong person.

(23) S J'ai venu te prendre.

(24) S Qu'est-ce que ça veut dire 'il fit le fanfaron'?
S Il moqueur.

T Il était moqueur.

(25) T Qu'est-ce que les esclaves faisaient?
S Elle le mangent.

(26) T As-tu l'intention...
S A-vous l'intention.
T Pardon?
S A-vous.
T A-vous?
S Avez-vous.

(27) T C'est moi qui vais la conduire dans le paturage.
S C'est nous qui vont les...
T Qui vont?
S X X X
T C'est nous qui allons.

Word order. Errors include utterances where the word order is clearly not French word order such as misplaced adverbials, adjective/noun sequences and mislocated pronouns.

(28) S Tu as dit déjà.

(29) S Un rouge chemise.

(30) S on mets les dans les bouilloires.

(31) S Est-ce que tu sais elle?

Reflexives. This category concerns the use of verbs non-reflexively that are reflexive.

(32) S C'est plus lourd donc ça prend plus de temps pour arrêter.
T Pour s'arrêter.

Pronunciation. These errors include obvious mispronunciations and un-made liaisons.

(33) S Il n'y a pas de messages pu/publi/
T Pu-bli-ci-taires. Tout le monde.
Ss Pu-bli-ci-taires.

(34) S (reading) Cette année fut si froid que les poissons gelave.
T Gelèrent.

(35) S Il était # une fois.
T Fais la liaison.

(36) S Des petits # avions.

The original tabulation of errors was carried out by a native speaker of French. A list of errors was made for each transcript, and each error was classified. A second researcher, a native speaker of English, checked the classification of errors. On the basis of the judgement of the second researcher, a small number of errors — on average, one per transcript — were reclassified. Thus, although a second, independent tabulation of the errors was not made, the classification of errors into categories was verified.

5:3 Results

The results from the error counts and classification are displayed in Tables 13 through 15. Table 13 (p. 262) shows the average number of student errors made in French across ten grade 6 immersion classes. The figures shown in Table 13 are necessarily lower than the average number of errors students make as they are based only on students' public talk in class, and of that, only on that which was audible enough to be transcribed. Even so, the average frequency of student error was 77.5 (per average of 106 minutes of recorded time). Table 13 indicates the percentage of the total number of errors that fall into each linguistic category. The figures appear to approximate the frequency of the linguistic categories in speech, with the highest percentages noted for gender (22.1%), article (17.2%) and verb (13.5%) errors and the lowest percentage for reflexive errors (3.4%). Another way of interpreting the error rate is that the higher frequencies represent the unavoidable categories (the use of nouns and verbs), while the lower frequencies represent the avoidable categories (use of reflexives and pronouns).

Table 14 (p. 262) shows the percentage of errors made by students which were corrected or left uncorrected. Overall 22.7% of student errors were corrected, while the remainder (77.3%) were ignored by the teachers. Table 15 indicates that when the overall percentage of errors corrected is separated into grammatical errors and pronunciation errors, 19% of grammatical errors and 66.7% of pronunciation errors were corrected. These figures are respectively lower and higher than those indicated by similar categories in Chaudron (1986). The mean percent across the studies summarized by Chaudron of grammatical correction is 51% with no study showing less than 36%; whereas the mean percent across the studies of phonological correction is 48% with the highest being the same as in the current study, 67%.

Table 14 also indicates the percentage of errors within each linguistic category that were corrected. Leaving aside pronunciation, it appears that there is a higher rate of correction among the categories with the highest frequency of error: gender (21.6%), articles (37.6%) and verbs (21%).

With these categories having the highest frequency of error, and with them being corrected the most frequently, it must seem to teachers as though they are continually involved in a battle to eradicate these errors. However, students seem almost impervious to these corrections. Consider, for example, the following situation (class No. 316). Students have been acting out a play, reading from a script. When they finish, the teacher asks the student for comments about the way in which the parts have been played. Over a ten-minute period, a number of comments are made. Interspersed in the ten minutes the following sequence occurs.

(37)
S J'ai pensé uhm qu'elle était très bonne parce que sa voix était très fort.
T Sa voix était très forte. C'est vrai, oui.

.
.
S	Je pense que c'était très bon parce qu'elle avait beaucoup d'expression dans son/son voix.
.
.
.
S	(Name) était excellent X X uhm voix très fort et beaucoup d'expression.
T	Oui, excellent, voix très forte, beaucoup d'expression.
.
.
.
S	euh sa voix était très euh
T	Forte.
S	Fort mais je pense qu'il peut avoir un peu plus de uhm d'expression.
T	Un peu plus d'expression, la voix très forte, oui.
.
.
.
S	X X bon voix.
T	Il a une très bonne voix, oui.
.
.
.
S	Je pensais que son voix était
T	Sa
S	Sa voix était très bien mais X X.

Just what grammatical 'message' are the students getting? Although the teacher provides students with feedback about the gender of *voix*, the message is a confusing one. In one case, no indication is given that the student is incorrect, even though the student is clearly hesitant: *son/son voix*. In several instances the teacher repeats the student, correcting the error as she does. However, in another instance, the teacher repeats the student correcting one error, but leaving another: *voix très forte*. If the message is that what the teacher repeats is correct, then the message is that *voix* does not need an article. (Using nouns without articles occurs frequently in the transcripts and accounts for the majority of article errors observed.) In another instance, the teacher provides the student with the correct adjective *forte*, which the student repeats as *fort*.

Consider another example. In one class (No. 326), during a grammar lesson, the teacher explains how to form the *plus-que-parfait*. She then continues by asking students to give an example using the auxiliary *être*.

(38)
S	J'étais vu.
T	J'étais vu, ça se fait oui.
S	J'avais mangé.
T	J'avais mangé. Un autre avec l'auxiliaire 'avoir'.
S	J'étais mangé.
T	Tu étais mangé, oui, oui, par un animal probablement.

The message here seems to be that être and avoir can be substituted for each other. The expansion in the last sentence is, however, not a neglible clue as the to meaning of the passive. Later though, in the same class, students make the following auxiliary errors; (39) was corrected, whereas (40) and (41) were not.

(39) S J'ai venu te prendre.
T Pardon?
= S Je suis venu.

(40) S J'étais très froid dehors alors j'ai revenu dans la maison.

(41) S J'ai allé en haut.

Obviously, these are but snippets of the feedback, or lack of it, that students in these classes receive. Nevertheless, they are illustrative of many such examples found in the transcripts of these classes. Their interest lies in the uninformative iand contradictory messages they impart. This lack of consistent and non-ambiguous feedback surely cannot aid learning.

There seems to be little sense in which students are 'pushed' towards a more coherent and accurate production of French. When they are corrected, that frequently suffices and there is no further follow-up. In relatively long student turns, teachers rarely make corrections at all. For example, in one class (No. 126), students were asked to summarize or read aloud what they had written about their favourite TV program. Over 65 errors were noted in this portion of the transcript. However, not one error was corrected.

In immersion classes the goal of language learning through content learning is paramount. This brings with it the questions of when to correct, what to correct and how to correct so well illustrated by examples (37) through (41). These are questions of major pedagogical import that the joint teaching of language and content must seriously begin to address. For further discussion on this point, see Appendix D, pp.317-341.

6. CONCLUDING COMMENT

Through an examination of what actually occurs in immersion classes, we are beginning to understand the language learning outcomes. Hypotheses based on these observations need to be explored through materials development and research projects carried out cooperatively with immersion teachers.

Footnotes

1 Transcription conventions were as follows. Normal orthography was used except for punctuation. Note, however, that '?' and '!' were used to convey rising and rising-falling intonation contours respectively. Series of periods correspond to pauses, so '..' is a short pause (approximately 1 to 2 seconds), '...' is a longer pause (approximately 3 seconds), and '....' is a very long pause (4 seconds or more). English pronunciations are put between single quotes. Slash bars '/' indicate interruptions, false starts or corrections. Overlapping speech is marked out by raised degree symbols '°text°' with pairs of one, two or even three degree symbols marking the beginning and ends of overlaps (depending upon the number of sequences of overlapping). Double quotes ("text") are used to indicate metalinguistic use of language, i.e. language talking about language. Capital letters correspond to focus stress, except for single letters when they indicate the beginning of an utterance. 'XX' is used to write down unintelligible utterances. Bold letters are used for metalinguistic highlighting. "T" refers to text produced by the teacher, while "S" refers to students. It should be noted that numbers following "S" are simply to encode the fact that the same individual is not speaking. They do not indicate that the speaker is always the same (i.e. they are not rigid designators in the logical sense). In the absence of videotapes, it was not possible to clearly identify persons.

2 Psychologists have provided the richest literature on the subject. Educators and linguists only now appear to be developing a comparable interest.

3 It should be emphasized that this speculation is put forth as an empirical hypothesis, one worthy of investigation and not because the author holds it to be true. In fact we suspect that specifically linguistic parameters will be brought to bear even if second language learning involves hypothesis testing of a sort which simply does not occur in the first language situation. These suspicions arise from observations concerning the absence of certain kinds of errors involving lexical items in second language production. Thus there is no evidence that L2 learners construct words on the basis of semantic primitives as opposed to structural ones like roots, stems and affixes.

4 No attempt was made to calculate how much time was devoted to a particular activity, nor how many times certain kinds of intervention occurred, consequently observations as to the generality of certain events or actions are entirely subjective.

5 That spontaneous inventions may not match actual words is a problem distinct from the creation of ungrammatical forms and teachers should be aware of the distinction between actual words sanctioned by use -- and appearing in dictionaries — and possible but non-existent words, on the one hand, and words which are not grammatical because they violate some rule of the grammar on the other hand. Thus *chienette* "little dog" or *chatette* "little cat" are possible but non-existent words constructed on the basis of a rule attaching the diminutive suffix *-ette* to a concrete noun stem (compare *fillette* "little girl", *jeunette* "young girl" etc.). The sanctioned, existing words, are, of course *chiot* "puppy" and *chaton* "kitten",

formed from the diminutives -ot and -on. In contrast, the following are simply ungrammatical meaning "a little, young or cute person who does X where "X" corresponds to the meaning of the verb stem, because the suffix -ette doesn't attach directly to verb stems: *mangette, *conduisette, *dormette. We will present an instance below where the teacher fails to make this distinction, and therefore fails to encourage a student for exploiting a rule of word formation while still presenting the actual word used.

6 An anecdote might be illustrative. A relative of one of the authors once claimed that she did not "speak" English before going to school. When it was pointed out that the person in question was not speaking Chinese or Russian at age three, the relative stopped, thought for a moment and then claimed that what she had meant was that the author had not spoken "real English" before entering school. Clearly, this relative had made a rather strong identification between written English and English as a whole. This confusion is by no means uncommon in literate societies.

7 The qualifications are required because comics, cartoons, plays and certain styles of speech-writing mimic natural speech. The extent of their success is a moot point.

8 This fact may result from the permanency of the written media and social taboos more strongly connected with formal situations.

9 These properties of English reflect directly certain historical realities, namely the conquest of England by the Normans, the ensuing bilingualism of the ruling classes, and the subsequent relexification of the language. For some discussion of historical change in English, see Lightfoot 1979.

10 It should be noted that word frequency lists, which often serve as the source for particular choices of vocabulary for inclusion in texts, may not be the best means of accomplishing this objective. A given lexical item. e.g. spatule or râpe might be common and necessary in a conversation conducted on how to prepare and cook some food and still be infrequent in conversations about food in general and even more infrequent in the language as a whole. Hence, if teachers want their students to be able to carry out certain functional activities and use language appropriate to those activities at the same time, it is vital to give some serious thought as to how one actually talks while carrying out the tasks.

11 Keeping in mind certain properties of language such as the greater or lesser collocatability of words, their domain relatedness, their frequency in discussions of certain topics, etc., as well as institutional facts such as ministry of education guidelines.

12 This, of course, does not imply that students will be taught the words they need to know to express their own thoughts since the vocabulary of readings is often theme-based (see van Ek and Alexander 1980, Thornton-Smith 1972).

13 Indeed anecdotal evidence suggests that French immersion teachers often perceive the texts available for use in their classes to be too hard. The texts in question are written for native speakers.

14 Following Barwise and Perry (1983) one can and should draw a distinction between these two aspects of a lexical item's semantics. Meaning represents conventional linguistic constraints on expressions whereas interpretation is derived from meaning, the discourse situation (speaker identity, speech act, locale and time), situation connections (to objects, properties, times and places), and resource situations (including speaker perceptions, commonplace knowledge, knowledge from previous discourse, shared assumptions about the discourse situation, etc., see Barwise and Perry 1983, pp. 32-36).

15 It might turn out to be true that learning new information about words, e.g. some morphological principle (see section 2:6), involves different processes than learning simply new items. Thus, it might be true that new items could be learned quite well alongside new activities if all that is required of the learner is the association of a sound and a meaning. This speculation merits further investigation.

16 But they can be the object of planned instruction and taught in a systematic fashion by, say, contrasting their idiosyncratic features with those of words which are part of a system and which are rule-governed. If it turns out to be true that rule-governed linguistic patterns are easier to learn than idiosyncratic linguistic information, even when the latter is presented according to some systematic fashion, e.g. mnemonic technique, then this fact (if fact it be) would lend support to claims about the specificity of linguistic generalizations. In other words, patterns arising out of the organization of the grammar would be more relevant to language learning than patterns arising out of non-linguistic systems such as the associative networks that mnemonic techniques typically exploit. Again, these speculations merit empirical study.

17 Thus people have a tendency to interpret mossy bank as a reference to a certain kind of land configuration and not a financial institution. For discussion of some of the relevant psycholinguistic literature on meaning and parsing, see Tanenhaus, Carlson and Seidenberg 1985.

18 The first kind of rule has been the focus of a considerable amount of research in various models of Natural Phonology (see Dressler 1984). The second kind of rule has been the focus of research in the Sound Pattern of English framework (Chomsky and Halle 1968), i.e. standard linear generative phonology, and more recently in lexical phonology (see Kaisse and Shaw 1985).

19 This is as true of first language learners as it is of L2 learners. Many Canadians have acquired the pronunciation /rəˈspəyt/ despite the fact that it is not sanctioned by either British or American dictionaries. It does appear as a natural pronunciation on the basis of the spelling. Similarly, a journalist appearing on national television was overheard saying /ˈatoˈmeytən/, instead of the more usual (and correct) /əˈtaməˈtən/. This error also appears to be caused by the spelling interfering with pronunciation rules.

20 This problem arises particularly in French which does not have any special orthographic marking to indicate that a word is a clitic. The students in our classes tended to treat preverbal atonic pronouns as stressable words, failing to

reduce their vowels and inserting pauses in between them and the verbs that followed. Similar comments can be made about their treatment of prenominal determiners.

21 There appears to be [illegible]ne serious confusion about this in the psycholinguistic literature. Thus, Karmiloff-Smith (1981) and Tucker, Lambert and Rigault (1977) refer to knowledge of suffixation and suffixes as phonological knowledge. Speaking formally, phonological units are those entities which are utilized as primitives and derived constituents of phonological representations. They are the meat of phonological generalizations. Morphological units such as suffixes, in contrast, are the necessary elements of morphological representations and encode quite different kinds of hypotheses about the organization of language. Thus, phonological units and morphological units do not constitute a natural class, although they may be instantiated by the same bits of noise.

22 Since the prefix is a unique morphological unit we follow convention and assign it one written form, despite the fact that it may have various pronunciations and some of those pronunciation variations may be recorded in spelling. Here the prefix is con although it has one spelling as com.

23 Lexical Redundancy rules allow one to relate lexical entries eliminating the explicit statement of redundant information. An example from English is given below

$$\begin{bmatrix} +\text{V} \\ +\ \text{NP}_1 \_\_ \\ \text{BREAK}\ (\text{NP}_1) \\ \quad\text{Theme} \end{bmatrix} \qquad \begin{bmatrix} +\text{V} \\ +\ \text{NP}_1 \_\_ \text{NP}_2 \\ \text{CAUSE}\ (\text{NP}_1 \quad \text{BREAK}\ (\text{NP}_2)\ ) \\ \quad\text{Agent} \qquad\qquad \text{Theme} \end{bmatrix}$$

The double-headed arrow says there is a regular relationship between the intransitive break and the transitive break, this relationship being causation. Since this relation holds true of numerous pairs of English verbs (forming a subset of the verbal lexicon), the lexical redundancy rule attempts to state that speakers know the nature of this subset -- its syntactic and semantic characteristics.

24 The general absence of forms [ɑ̃tmɑ̃] may well be due to a phonological restriction preventing a word from ending with two identical rhymes of this shape. Where an alternative form is available which does not infringe this restriction, adverbs appear, hence consciemment, but not *conscientement, puissamment but not *puissantement. Grevisse notes some exceptions to this constraint -- lentement présentement, véhémentement.

25 It should be apparent that modifying expressions can denote more than one element in a collection, e.g. plusieurs, deux, nombreux, etc. Hence the ill-formedness of *Un homme nombreux est venu. But morphological plurality and semantic plurality are separate phenomena. Consider in this regard The crowd was restless where the noun denotes a plurality but is grammatically singular. See also the examples in the text.

26 The explanation for the plurality of the words in both cases may be that there is a distinct word which is semantically-related and historically older which occurs in the singular.

27 The distinctions can be seen as part of a larger problem of developing awareness of dialect differences, and making students aware of common attitudes towards different speech communities.

28 The converse is not true; planned activities are not necessarily restricted to the written mode.

29 These difficulties arising in part because of the considerable degree of homonymy displayed by various words, those that "agree" and those that don't.

30 These superscripts indicate that speakers are talking simultaneously.

31 There was also vocabulary teaching in the narrower, more common sense since within the grammar lesson students learned some linguistic terminology (e.g. "A *subject* is an expression ...").

32 In linguistic terms one would say that these subjects lack a semantic role; in the lesson in question students were taught (on the basis of information in the text) that *il* corresponds to the *neuter* gender whenever it cannot be replaced by a noun. The grammatical account is questionable but the facts remain - *il* cannot be replaced by a noun with weather verbs.

33 *Deverbal* nouns are derived by rule from verbs while *denominal* verbs are derived by rule from nouns.

34 With enough parts to the question to separate out the different kinds of secondary cases (and their distinct meanings).

35 This is because it can convey other meanings totally unrelated to the concept of "small size".

36 It may also be a regional marker. For example, in Quebec, university students may use *tu* or *toi* with their professor. Such usage is uncommon in Europe.

37 There is some question as to whether it is 'appropriate' to address a teacher as *vous*. In the two francophone classes for which we have tape recordings, the grade 6 students addressed their teacher as *tu*. They have also called their teachers by their first names, not by *madame* or *monsieur*. However, in discussing this issue with teachers in a Franco-Ontarian school, they each said that they insisted on the use of *vous* when their students addressed them. What does seem clear from our native French-speaking informants on the project is that it seems sociolinguistically inappropriate to address someone as *madame* and *monsieur* AND use *tu*.

[38] We recognize the limitations of this kind of classification, but given the data, few alternatives existed.

[39] We infer error in (5) because we assume that the article is the gender indicator. The teacher, we assume, takes this view. However, the student's problem might be in articulating /œ̃/ vs /yn/. Similarly, (6) could be a co-reference problem and (8) an agreement error.

[40] These three are quite different in nature. (9) and (10) relate to the structure of NP's, while (11) is a discourse error.

References

Abberton, Evelyn. Some persistent English vocabulary problems for speakers of Serbo-Croatian. **English Language Teaching**, 1968, **22**, 167-172.

Albert, Martin and Obler, Lorraine. **The Bilingual Brain, Neuropsychological and Neurolinguistic Aspects of Bilingualism**. New York: Academic Press, 1978.

Anderson, John R, **Cognitive Psychology and Its Implication**, 2nd. edition. New York: W.H. Freeman and Company, 1980.

Aronoff, Mark. **Word Formation in Generative Grammar**. Linguistic Inquiry Monograph 1. Cambridge, Mass.: MIT Press, 1976.

Barwise, Jon and Perry, John. **Situations and Attitudes**. Cambridge, Mass.: MIT Press, 1983.

Blum-Kulka, Shoshana and Levenston, Eddie. Universals of lexical simplification. In C. Faerch and G. Kasper (eds.) **Strategies in Interlanguage Communication**. London: Longmans, 1983, 119-139.

Butterworth, Brian. Lexical representations. In B. Butterworth (ed.) **Language Production II**. New York: Academic Press, 1983, 257-294.

Canale, Michael. From communicative competence to communicative language pedagogy. In J. Richards and R. Schmidt (eds.), **Language and Communication**. London: Longman, 1983, 2-27.

Canale, Michael and Swain, Merrill. Theoretical bases of communicative approaches to second language teaching and testing. **Applied Linguistics**, 1980, 1, 1, 1-47.

Carrell, Patricia. Cohesion is not coherence. Paper presented to the Second Language Research Forum, Los Angeles, April 1982.

Chau, Tran Thi. The Concept of difficulty in second-language learning and teaching. Unpublished doctoral dissertation, University of Toronto, 1972.

Chaudron, Craig. A descriptive model of discourse in the corrective treatment of learners' errors. **Language Learning**, 1977a, **17**, 29-46.

Chaudron, Craig. Teachers' priorities in correcting learners' errors in French immersion classes. **Working Papers in Bilingualism**, 1977b, **12**, 22-44.

Chaudron, Craig. The role of error correction in second language teaching. **Working Papers**, 5(2), Dept. of ESL, Univ. of Hawaii at Manoa, 1986, 43-81.

Chomsky, Noam and Halle, Morris. **The Sound Pattern of English**. New York: Harper & Row, 1968.

Courchêne, R. The error analysis hypothesis, the contrastive analysis hypothesis, and the correction of error in the second language classroom. **TESL Talk**, 1980, 11(2), 3-13; 11(3), 10-29.

Dell, François. **Les Règles et les sons.** Collection Savoir. Paris: Hermann, 1973.

Dell, François and Selkirk, Elisabeth. On a morphologically-governed vowel alternation in French. In Keyser (ed.) **Recent Transformational Studies in European Languages.** Cambridge, Mass.: MIT Press, 1978, 1-51.

Dressler, W. Explaining natural phonology. In Ewan and Anderson (eds.) **Phonology Yearbook 1**, 1984, 29-52.

Fanselow, John. The treatment of error in oral work. **Foreign Language Annals**, 1977, **10**, 583-593.

Fowler, Carol, Napps, Shirley and Feldman, Laurie. Relations among regular and irregular morphologically-related words in the lexicon as revealed by repetition priming. **Memory and Cognition**, 1985, **13**, 3, 241-255.

Fuld, Paula A. and Buschke, Herman. Stages of retrieval in verbal learning. **Journal of Verbal Learning and Verbal Behavior**, 1976, **15**, 401-410.

Galisson, Robert. **Lexicologie et enseignement des langues.** Paris: Hachette, 1979.

Genouvrier, E. and Poulin, D. **De la Langue parlée à la langue écrite.** Montréal: Les Éditions Françaises, 1973.

Gregory, Michael and Carroll, Susanne. **Language and Situation: Language Varieties and their Social Contexts.** London: Routledge and Kegan Paul, 1978.

Grice, H.P. Logic and conversation. William James Lectures, Harvard. In Cole and Morgan (eds.) **Syntax and Semantics 3; Speech Acts.** New York: Academic Press, 1967, 41-58.

Hammar, Petra. Cognates as an efficient approach to second language acquisition. Unpublished doctoral dissertation, University of Alberta, 1978.

Harley, B. **Age in Second Language Acquisition.** Clevedon, England: Multilingual Matters, 1986.

Harley, B. and Swain, M. An analysis of the verb system used by young learners of French. **Interlanguage Studies Bulletin**, 1978, **3**(1), 35-79.

Harley, B. and Swain, M. The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper and A. Howatt (eds.), **Interlanguage.** Edinburgh: Edinburgh University Press, 1984, 291-312.

Higa, Masanori. The psycholinguistic concept of "difficulty" and the teaching of foreign language vocabulary. **Language Learning**, 1965, **15**, 167-179.

Jackendoff, Ray. **Semantic Interpretation in Generative Grammar.** Cambridge, Mass.: MIT Press, 1972.

Kaisse, Ellen, and Shaw, Patricia. Phonology and the lexicon: on the theory of lexical phonology. In Ewen and Anderson (eds.) **Phonology Yearbook 2**, 1985.

Karmiloff-Smith, Annette. **A Functional Approach to Child Language: a Study of Determiners and Reference.** Cambridge Studies in Linguistics 24. Cambridge: Cambridge University Press, 1981.

Kenyeres, A. Comment une petite hongroise apprend le français. **Archives de Psychologie**, 1938, **26**, 321-366.

Levenston, Eddie and Blum, Shoshana. Aspects of lexical simplification in the speech and writing of advanced adult learners. In S.P. Corder and E. Roulet (eds.) **The Notions of Simplification, Interlanguages and Pidgins and their Relation to Second Language Pedagogy.** Geneva: Droz, 1977.

Leopold, Werner. **Speech Development of a Bilingual Child. Vocabulary, Vol. 1.** Evanston, Ill.: Northwestern University Press, 1939.

Lieber, Rochelle. On the organization of the lexicon. Unpublished doctoral dissertation, MIT, Cambridge, Mass., 1980.

Lightbown, Patricia. Can language acquisition be altered by instruction? In K. Hyltenstam and M. Pienemann (eds.) **Modelling and Assessing in Second Language Acquisition.** Clevedon, Avon: Multilingual Matters 18, 101-112.

Lightfoot, David. **Principles of Diachronic Syntax.** Cambridge: Cambridge University Press, 1979.

Long, M. Input, interaction, and second language acquisition. Unpublished Ph.D. dissertation, University of California, Los Angeles, 1980.

Lucas, E. Teachers' reacting moves following errors made by pupils in post-primary English as a second language class in Israel. MA thesis, School of Education, Tel Aviv University, 1975.

Meara, Paul. Vocabulary acquisition: A neglected aspect of language learning. **Language Teaching and Linguistics Abstracts**, 1981, **13**, 4 (Oct.), 221-246.

Meara, Paul. The study of lexis in interlanguage. In A. Davies, C. Criper and A.P.R. Howatt (eds.) **Interlanguage**, Edinburgh: University of Edinburgh, 1984, 225-233.

Meara, Paul. The classical research in vocabulary acquisition. In S. Gass (ed.) **Studies in Second Language Acquisition**, (in press), **9**, 2.

Napps, S. Morphological, semantic and formal relations among words and the organization of the mental lexicon. Unpublished doctoral dissertation, the University of Dartmouth, Dartmouth, New Hampshire, 1985.

Pica, T. and Doughty, C. Input and interaction in the communicative language classroom: a comparison of teacher-fronted and group activities. In S. Gass and C. Madden (eds.) **Input in Second Language Acquisition.** Rowley, Mass.: Newbury House, 1985, 115-132.

Pylyshyn, Zenon. **Computation and Cognition: Toward a Foundation for Cognitive Science.** Cambridge, Mass.: Bradbook Books, MIT Press, 1985.

Rogers, Theodore. On measuring vocabulary difficulty: an analysis of item variables in learning Russian English vocabulary pairs. **IRAL**, 1969, **7**, 327-343.

Ronjat, J. **Le Développement du langage observé chez un enfant bilingue.** Paris: H. Champion Librairie Ancienne, 1913.

Salicia, C. Testing a model of corrective discourse. MA in TESL thesis, UCLA, 1981.

Selkirk, Elisabeth. The role of prosodic categories in English word stress. **Linguistic Inquiry**, 1980, 563-605.

Selkirk, Elisabeth. **The Syntax of Words.** Cambridge, Mass.: MIT Press, 1982.

Snow, Catherine. Mother's speech to children learning language. In **Child Development**, 1972a, **43**, 459-565.

Snow, Catherine. Young children's responses to adult sentences of varying complexity. Paper presented to IVth Internation Congress of Applied Linguistics, Cpenhagen, 1972b.

Snow, Catherine and Ferguson, C.A. (eds.) **Talking to Children: language input and acquisition.** Cambridge: Cambridge University Press, 1977.

Stanners, Robert F., Neiser, James J., Hernon, William P., and Hall, Roger. Memory representation for morphologically-related words. In **Journal of Verbal Learning and Verbal Behaviour**, 1979, **18**, 399-412.

Stemburger, J. The lexicon in a model of language production. Unpublished doctoral dissertation, University of California, San Diego, 1982.

Swain, M. Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden (eds.), **Input in second language acquisition.** Rowley, Mass.: Newbury House, 1985, 235-253.

Swain, M. and Lapkin, S. **Evaluating Bilingual Education: A Canadian Case Study.** Clevedon, England: Multilingual Matters, 1982.

Tanenhaus, Michael K., Carlson, Greg N., and Seidenberg, Mark S. Do listeners compute linguistic representations? In Dowty, Karttunen and Zwicky (eds.) **Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives.** Cambridge: Cambridge University Press, 1985, 359-408.

Thornton-Smith, C.B. **Graded French Vocabulary Topics.** London: Oxford University Press, 1972.

Tranel, Bernard. **Concreteness in Generative Phonology: Evidence from French.** Berkeley, California: University of California Press, 1981.

Tucker, G. Richard, Lambert, Wallace E. and Rigault, André. **The French Speaker's Skill with Grammatical Gender: An Example of Rule-governed Behaviour**. The Hague: Mouton, 1977.

Varonis, E.M. and Gass, S. Non-native/non-native conversations: A model for negotiation of meaning. **Applied Linguistics**, 1985, 6, 71-90.

Van Ek, J.A. and Alexander, L.G. **Threshold Level English**. Oxford: Pergamon Press, 1980.

West, M. Vocabulary selection and the minimum adequate vocabulary. **English Language Teaching**, 1954, 8, 4.

Yoshida, Midori. The acquisition of English vocabulary by a Japanese-speaking child. In E. Hatch (ed.) **Second Language Acquisition: A Book of Readings**. Rowley, Mass.: Newbury House, 1978.

Table 1: Instructional Activities Schema

INSTRUCTIONAL ACTIVITIES	
planned	unplanned
systematic	haphazard
written input	oral input
L1 prior knowledge	L2 prior knowledge
direct control of item choice	indirect control
interpretation	formal focus

Table 2: Aspects of Vocabulary Learning

ASPECTS OF VOCABULARY LEARNING	
attending to data	not attending
systematic	idiosyncratic
iconic	echoic memory
self-	teacher initiated

Table 3: Linguistic Aspects for Learning

LINGUISTIC ASPECTS FOR LEARNING		
COMPONENT	RULE-GOVERNED	IDIOSYNCRATIC
phonology	features, units	underlying features & forms
morphology	word formation rules	exceptional features
syntax	categories from rules	class/subcategorization
semantics	scopal properties	roles
"graphology"	sound-letter correspondances	irregularities
sociolinguistics	standard properties	dialect/register restrictions
discourse	coherence/cohesion	?

Table 4

Average Frequency and Percent Use of Singular, Plural and Generic Functions of tu and vous by Ten Grade 6 Immersion Teachers

	"tu"		"vous"	
	%	($\bar{X}$ freq)	%	($\bar{X}$ freq)
singular	46.0	(78.0)	0.9	(1.6)
plural	3.7	(6.3)	44.5	(75.4)
generic	3.0	(5.0)	1.9	(3.2)
total	52.7	(89.3)	47.3	(80.2)

Table 5

Average Frequency and Percent Use of Informal, Formal and Other Functions of tu and vous by Ten Grade 6 Immersion Teachers

	"tu"		"vous"	
	%	($\bar{X}$ freq)	%	($\bar{X}$ freq)
informal	49.7	(84.3)	NA	
formal	NA		0.1	(0.2)
other	3.0	(5.0)	47.2	(80.0)
total	52.7	(89.3)	47.3	(80.2)

Table 6

Average Frequency and Percent Use of Singular, Plural and Generic Functions of tu and vous in the Public Talk of Grade 6 Students in Ten Immersion Classes

	"tu"		"vous"	
	%	($\bar{X}$ freq)	%	($\bar{X}$ freq)
singular	55.6	(12.5)	3.6	(0.8)
plural	0.9	(0.2)	0.4	(0.1)
generic	39.5	(8.9)	0.0	(0.0)
total	96.0	(21.6)	4.0	(0.9)

Table 7

Average Frequency and Percent Use of Informal, Formal and Other Functions of *tu* and *vous* in the Public Talk of Grade 6 Students in Ten Immersion Classes

	"tu"		"vous"	
	%	($\bar{X}$ freq)	%	($\bar{X}$ freq)
informal	12.9	(2.9)	NA	
formal	43.6	(9.8)*	2.7	(0.6)
generic	39.5	(8.9)	1.3	(0.3)
total	96.0	(21.6)	4.0	(0.9)

* These represent examples of students using *tu* to address the teacher

Table 8
Category, Frequency and Percent of Student Talk in Teacher-fronted Activities in 90 Minutes of The French Portion of Nine Grade 3 Immersion Classes

	Student-Initiation						Teacher Initiation																					
							Pre-Planned				Un-Planned																	
	Self			Other Student			Read aloud (from own production)			Spell	Repeat			Complete			Select from limited choice			Find own words			Read aloud from text			Recite from memory		
Class	M	P	E*	M	P	E	M	P	E	M	M	P	E	M	P	E	M	P	E	M	P	E	M	P	E	M	P	E
333	2	10	51	3	1	1	0	0	0	4	33	6	5	3	1	0	105	5	29	4	22	64	0	0	12	0	0	0
313	0	2	42	4	0	2	0	0	3	0	0	0	1	2	1	0	87	7	33	1	0	14	1	0	0	0	0	0
323	0	0	14	0	0	1	0	0	29	0	0	0	1	0	0	0	19	1	5	0	1	32	0	0	0	0	0	0
353	3	2	21	1	0	0	0	0	0	0	9	5	0	0	0	0	53	25	18	14	22	58	0	0	6	0	0	0
223	9	0	50	0	0	17	0	0	0	0	13	9	3	1	0	0	45	3	35	1	1	17	0	0	24	0	0	1
243	0	0	8	11	6	1	0	0	4	0	24	27	20	11	1	0	34	26	52	3	6	45	0	0	16	0	0	0
343	6	2	31	3	3	5	3	1	0	7	6	2	0	1	0	0	72	2	1	0	0	14	0	0	15	0	0	0
233	3	2	67	11	0	6	0	0	0	0	22	10	3	0	0	0	44	3	17	0	3	14	0	0	12	0	0	0
213	2	1	26	7	2	1	16	0	0	23	17	9	0	8	0	0	12	2	14	16	0	12	0	12	13	0	0	0
Mn%	1.3	0.8	16.0	1.9	0.6	1.8	1.1	0.1	3.4	1.9	5.5	3.0	1.4	1.2	0.1	0.0	23.6	3.3	9.5	1.9	2.2	13.7	0.1	0.7	4.7	0.0	0.0	0.05
%	18.1			4.3			4.6			1.9	9.9			1.3			36.5			17.9			5.5			0.5		

% 22.4 Student-Initiation

77.6 Teacher-Initiation

---

*M = Minimal - one or two words
P = Nominal, adverbial or verb phrase
E = Extended - one or or more clauses

Table 9

Type, Frequency and Percent of Student Talk in Category "Selecting from Limited Choice" in Teacher-fronted Activities in 90 Minutes of the French Portion of Nine Grade 3 Immersion Classes

Class	Yes/No	Responses semantically controlled			Responses syntactically controlled		
		M	P	E*	M	P	E*
333	41	1	5	19	63	0	10
313	36	51	7	33	0	0	0
323	19	0	1	5	0	0	0
353	16	0	25	16	37	0	2
223	25	15	3	13	5	0	22
243	12	19	16	24	3	10	28
343	22	0	2	1	50	0	0
233	30	5	2	7	9	1	10
213	5	1	2	1	6	0	13
Mean	22.9	10.2	7.0	13.2	19.2	1.2	9.4
Mean %	31.7	9.7	7.8	14.4	21.6	1.2	13.7
%	31.7	31.8			36.5		

*M = Minimal - one or two words
P = Nominal, adverbial or verb phrase
E = Extended - one or more clauses

Table 10
Category, Frequency and Percent of Student Talk in Teacher-fronted Activities in 90 Minutes of The French Portion of Ten Grade 6 Immersion Classes

Student-Initiation							Teacher Initiation																					
							Pre-Planned				Un-Planned																	
	Self			Other Student			Read aloud (from own production)			Spell	Repeat			Complete			Select from limited choice			Find own words			Read aloud from text			Recite from memory		
Class	M	P	E*	M	P	E	M	P	E	M	M	P	E	M	P	E	M	P	E	M	P	E	M	P	E	M	P	E
116	4	1	13	0	0	0	16	3	2	8	5	9	1	6	2	0	71	19	22	6	2	13	0	4	13	0	0	0
146	7	6	27	11	1	2	0	0	0	0	4	3	0	9	1	3	127	34	18	7	10	25	0	0	1	0	0	0
326	0	1	30	6	1	1	0	0	3	56	10	7	5	1	0	0	58	18	34	1	6	41	1	10	56	0	0	5
336	3	0	9	2	0	5	0	0	1	0	5	5	3	0	0	0	25	5	5	8	10	69	0	0	0	0	0	0
156	2	3	45	8	0	12	0	0	0	0	1	0	4	3	0	0	33	12	40	1	1	13	0	2	19	0	0	0
356	4	9	69	4	1	5	0	12	8	21	5	1	0	1	0	0	59	3	10	5	9	87	0	0	0	0	0	0
316	0	1	15	1	0	0	0	0	0	4	0	0	0	0	0	0	24	4	12	14	9	83	0	0	14	0	0	0
346	2	0	5	0	0	0	0	0	0	0	15	3	0	0	0	0	21	3	17	1	0	14	5	1	17	0	0	0
126	1	4	15	1	0	1	0	0	15	2	4	0	0	0	0	0	21	5	3	1	0	14	0	0	3	0	0	0
136	5	6	54	4	4	3	5	15	3	25	5	0	0	3	0	0	180	10	2	6	1	13	3	1	10	0	0	0
Mn%	1.3	1.4	11.9	1.5	0.2	1.4	0.9	0.9	2.2	3.8	3.2	1.4	0.6	0.9	0.1	0.1	25.0	4.9	7.9	2.4	2.1	18.3	0.6	0.7	6.2	0.0	0.0	0.1
%	14.6			3.1			4.1			3.8	5.1			1.1			37.8			22.9			7.5			0.1		

% 17.6
Student-Initiation

82.4
Teacher-Initiation

---

*M = Minimal - one or two words
P = Nominal, adverbial or verb phrase
E = Extended - one or or more clauses

Table 11

Type, Frequency and Percent of Student Talk in Category "Selecting from Limited Choice" in Teacher-fronted Activities in 90 Minutes of the French Portion of Ten Grade 6 Immersion Classes

Class	Yes/No	Responses semantically controlled			Responses syntactically controlled		
		M	P	E*	M	P	E*
116	14	18	12	12	39	7	10
146	54	54	24	16	19	10	2
326	19	33	11	32	6	7	2
336	22	3	5	5	0	0	0
156	13	8	5	24	12	7	16
356	36	14	3	10	9	0	0
316	15	3	3	7	6	1	5
346	7	13	3	17	1	0	0
126	17	2	0	0	2	5	3
136	21	131	4	0	28	6	2
Mean	21.8	27.9	7.0	12.3	12.2	4.3	4.0
Mean %	31.2	22.8	7.5	16.4	11.6	4.9	5.5
%	31.2	46.7			22.0		

*M = Minimal - one or two words
P = Nominal, adverbial or verb phrase
E = Extended - one or more clauses

Table 12

Length of Student Talk in French for Grade 3 and 6 and in English for Grade 6

Grade	Language	Total No. of Utterances	Minimal (% of total)	Phrase (% of total)	Clause (% of total)	Sustained (% of total)	Sustained minus that read from text (% of total)
		$\bar{X}$	$\bar{X}$	$\bar{X}$	$\bar{X}$	$\bar{X}$	$\bar{X}$
3	French[a]	223.2	38.5	10.9	34.4	16.2	12.9
6	French[a]	226.7	39.5	11.7	30.1	18.7	14.9
6	English[b]	358.4	35.1	10.6	34.9	19.4	16.7

Minimal:	One or two words (includes spelling)
Phrase:	A nominal, adverbial or verb phrase
Clause:	One clause
Sustained:	More than one clause

a - based on 90 minutes of tape per class
b - based on an average of 88 minutes of tape per class

Table 13

Average Number of Errors in French Made by Grade 6 Students in their Public Talk in Ten Grade 6 Immersion Classes*

	gender	articles	contractions	pronouns	prepositions	verbs	word order	reflexives	pronunciation	other	total
$\bar{X}$	17.1	13.3	4.9	4.0	6.1	10.5	7.4	2.6	6.0	5.6	77.5
% of total	22.1	17.2	6.3	5.2	7.9	13.5	9.5	3.4	7.7	7.2	100

* based on average of 106 minutes of recorded time

Table 14

Percentage of Student Errors in French Corrected and Not Corrected in Each Linguistic Category in Ten Grade 6 Immersion Classes*

	gender	articles	contractions	pronouns	prepositions	verbs	word order	reflexives	pronunciation	other	total
% errors corrected	21.6	37.6	12.2	12.5	6.6	21.0	2.7	3.8	66.7	16.1	22.7
% errors not corrected	78.4	62.4	87.8	87.5	93.4	79.0	97.3	96.2	33.3	83.9	77.3

* based on average of 106 minutes of recorded time

Table 15

Percentage of Student Pronunciation and Grammatical Errors in French Corrected and Not Corrected in Ten Grade 6 Immersion Classes*

	grammar	pronunciation
% of errors corrected	19.0	66.7
% of errors not corrected	81.0	33.3

*based on average of 106 minutes of recorded time

Appendix A

Second Language Proficiency and Classroom Treatment in Early French Immersion

Birgit Harley

Ontario Institute for Studies in Education

Paper presented at the FIPLV/Eurocentres Symposium on Error in Foreign Language Learning: Analysis and Treatment, Goldsmiths' College, University of London,
September 2-6, 1985.

In response to public demand, French immersion is offered as an option to English-speaking students in increasing numbers of schools across Canada. Recent statistics (Commissioner of Official Languages, 1984) show close to 150,000 students enrolled in immersion programs of various kinds —'total' or 'partial' depending on the proportion of time devoted to schooling in French in the initial years of the program; 'early', 'middle', or 'late' depending on the starting point of the program.

The concept underlying immersion is that experience in using the second language for genuine communication is an essential aspect of the language learning process. Thus in an immersion classroom, français is not just a subject but also serves as the medium via which other school subjects, such as mathematics, social studies, art, music, science or physical education, are taught.

Of the program variations mentioned above, the most popular is that of 'early total' immersion beginning in kindergarten, and it is with this type of program that the present paper is concerned. Children in an early total immersion program receive their early schooling entirely in French, usually from teachers who are native or near-native speakers of French, and are gradually introduced to English language instruction as they progress through the elementary grades. By grade 5 or 6 (age 10-11), their instruction is divided about half and half between the two languages, with some school subjects taught in French and others in English. The program is maintained on a reduced basis into the high school years with, typically, a class period devoted to French language arts, and one or two other subjects offered in French (for further details, see Stern 1984; Swain and Lapkin 1982).

The purpose of this paper is to describe an ongoing series of studies concerned with the development of French proficiency by early total French immersion students.[1] Investigation of the students' proficiency in grades 5 and 6 is aimed at identifying both 'strengths' and 'weaknesses' in the second language, defined stringently in relation to native French-speaking norms. The analysis of errors plays an important role in this research, but it is not the only means of diagnosis. Similarly, in investigating classroom practices, the concern is not only with the treatment of overt errors but with the way in which problem areas in the second language are dealt with in general. Three main issues are considered here: 1. What are the characteristics of immersion students' French after several years in this innovative type of language teaching program? 2. How are problem areas currently being treated in the classroom context? 3. What, if any, additional treatment might be indicated? This research, conducted at the Modern Language Centre of the Ontario Institute for Studies in Education, can be roughly divided into the following stages:

(a) Analysis of the French proficiency of early total immersion students at the grade 5-6 levels

(b) Development of hypotheses to account for proficiency results

(c) Classroom observation to refine hypotheses

(d) Development of materials in accordance with hypotheses

(e) Experimental testing of materials to determine their impact on French proficiency (this stage of the research is about to take place).

In what follows, each of these stages is considered in turn.

(a) Analysis of the French proficiency of early total immersion students in grades 5-6

Evaluation of the students' French proficiency after six to seven years of the immersion experience has taken a variety of forms, including global comprehension tasks, cloze and multiple choice tests, written compositions and oral interviews designed to provide contexts for the expression of particular functions. Some of these tasks have been designed to focus specifically on grammatical, discourse, or sociolinguistic aspects of the second language (see e.g. Allen et al. 1983, Swain 1985). As an extremely rigorous, but diagnostically valuable, criterion against which to assess the proficiency of the immersion students, comparison data on the same tasks have been collected from native French speakers of the same age in the province of Quebec.

Discourse competence

Looking first at the strengths of the immersion students relative to native speakers, we find that they do exceptionally well on tasks involving global comprehension of discourse in French. By grade 8, for example, early total immersion students have obtained scores equivalent to those of native speakers on a listening comprehension test involving multiple choice questions based on taped extracts from, for example, radio broadcasts (Swain and Lapkin 1982).[2] This well-developed discourse competence extends to tasks designed specifically to tap the ability to interpret and produce logically coherent discourse (Allen et al. 1983). Grade 6 immersion students have, for example, received discourse coherence ratings for the production of written discourse that are similar to those obtained by native speakers, and for oral discourse that are only slightly lower than those of a native speaker comparison group. The ratings were awarded for coherence of discourse features such as reference to characters, objects and locations, and temporal and logical sequence (Allen et al. 1983). It is important to note that on such measures, even native-speakers in grade 6 did not achieve perfect scores. On the compositions, for example, the average discourse ratings for 69 immersion students and 23 native speakers were both at 1.5 out of a maximum of 2.0. A native-speaker comparison group serves a useful purpose here in tempering our sometimes unrealistically high expectations for the second language learner.

Grammatical competence

On measures of grammatical accuracy, as opposed to discourse coherence, it is clear that immersion students by grades 5 and 6 do not reach native-like competence, although they far outshine students in more traditional French-as-a-second-language programs (Swain and Lapkin 1982). Our assessment of immersion students' grammatical skills at these grade levels leads to the conclusion that although they are able to interpret and produce many grammatically accurate forms, they are prone to errors that clearly distinguish them from native speakers of their own age. Whether the language task consists of an oral interview, a discrete point multiple choice grammar test, a cloze test, or a written composition, the immersion students do not attain native-like levels of accuracy when scored, for example, for syntax, prepositions, verb forms, or gender (e.g. Allen et al. 1983; Harley 1979; Harley and Swain 1978; Swain and Lapkin 1982). In a conversational interview, for example, designed to elicit the use of verb forms in the future, the imparfait, the passé composé and the conditional, 69 grade 6 immersion students received an average accuracy score of 57 per cent compared with the close to 100 per cent score obtained by a comparison group of 23 native speakers (Allen et al. 1983).

Even in the grammatical domain, however, we find that native speakers do not always turn in a faultless performance. When compositions were scored for grammatical accuracy, for example, it was found that in one respect -- the spelling of homophonous verb inflections (e.g. -er, -é, -ez) -- grade 6 native speakers and immersion students had a similar average error rate (Allen et al. 1983). On a discrete-point multiple choice grammar test, native speakers averaged 80 per cent correct vis-à-vis an immersion score of 60 per cent. These less-than-perfect native speaker scores indicate that at least some of the grammatical errors made by these 10-11 year old second language learners have as much to do with their level of maturity as with their second language competence per se.

Accuracy scores are not the only useful diagnostic comparison between immersion students and native speakers. By comparing the use of forms to realize notions or functions required by particular tasks, we may find systematic differences between the second language learners and the native speakers that do not necessarily reveal themselves as overt errors. In assessing grade 6 immersion students' accuracy in the use of prepositions in written compositions, for example, it was noted that it was sometimes difficult to determine whether the error lay in the choice of preposition or the preceding verb, or indeed whether any error had actually been committed.[3] One of the composition topics involved the rescue of a kitten chased up a tree, and in telling their stories, students created numerous contexts for directional expressions (dogs chasing the kitten off the verandah, the kitten running away and up a tree, firemen bringing the cat down, owners taking it into their house, and so on). A detailed analysis of how direction was expressed in this narrative by 23 immersion students and 23 native speakers (Harley in press) revealed that the immersion students were making considerably less use of French verbs that combined motion and direction (e.g. ARRIVER, PARTIR, MONTER, RENTRER) than the native speakers, preferring instead a verb such as COURIR which combines motion and manner, and attempting to mark direction in a preposition phrase. For example: _trois chiens couraient jusqu'au balcon_; _Le petit chat coura dans un arbre_. Only by relating form with function and by making the comparison with native speakers did this frequent pattern of immersion language use become clear, a pattern doubtless influenced by the English mother tongue, but not necessarily manifested in errors that could be defined in any specific category (see also Harley and Swain 1984).

An analysis of the distribution of particular forms in the interlanguage of immersion students (Harley and Swain 1984) serves to demonstrate that native-like use of a target-language structure in a particular context is no guarantee that a learner is operating with the same underlying rule as a native speaker. Apparently correct use in grade 1 of a structure such as j'ai oublié may mean to the learner not 'I have forgotten' but 'I forget', il est tombé may mean not 'he has fallen' but 'he is falling', j'ai six ans may be analysed literally as 'I am six years'. At least some of the verb errors noted by grade among older students (e.g. *j'ai a oublié, *on a allé, *je suis dix ans) are thus not necessarily indications of 'backsliding' as dismayed teachers are apt to believe, but signs that the students have made progress in segmenting pronoun and auxiliary and sorting out the forms of ETRE and AVOIR. A similar kind of superficially correct performance without 'underlying competence' is manifested at the grade 5 and 6 levels in the use of the imparfait. Use of correct forms of the imparfait with inherently durative, non-dynamic verbs such as ETRE, AVOIR, VOULOIR, AIMER may not mean that the students have a grasp of the aspectual function of this verb form. In a detailed interlanguage analysis of oral interviews with five grade 5 immersion students (Harley and Swain 1978), for example, use of the imparfait was found to be restricted almost exclusively to the stative verbs ETRE and AVOIR, while habitual past actions or past actions in progress were usually expressed in the passé composé. An examination of the narrative compositions produced by grade 6 immersion students indicates that the imparfait may be used in the context of dynamic verbs in the written mode, but at least at times without apparent understanding of function. Provided with the following story opening: "Ce jour-là, comme d'habitude, la banque était pleine de monde. Tout d'un coup, trois bandits...", one student carries on as follows: "...entrait dans la porte. Ils prennait des fusils de leurs longues manteaux et ont demandé tout l'argent de la banque," using the imparfait inappropriately along with the passé composé to indicate the sequence of events.

In general, our analysis of the grammatical competence of grade 5 and 6 immersion students indicates that they tend to make errors in those relatively redundant rules of grammar that do not bear a heavy communicative load (e.g. number agreement rules, gender with respect to inanimate nouns, use of auxiliary ETRE with certain verbs in the passé composé), and that there is a continuing influence of the mother tongue, even when this is not manifested in overt errors.

Methodologically speaking, in assessing the grammatical competence of grade 5 and 6 immersion students, and the amount of progress they have made, we have found it necessary to keep in perspective not only error rates in particular grammatical categories but also the nature of the task assigned, the performance of native speakers on the same tasks, the relationship betweeen the use of second language forms and the functions they express (for the learners and the native speakers), and information about previous stages of development in the interlanguage of the students.

Sociolinguistic competence

The information we have about the 'sociolinguistic competence' of immersion students in grade 6, narrowly defined as the ability to adjust register in accordance with the social demands of a situation (Allen et al. 1983), is based largely on role-playing tasks (oral and written). The students are faced with social situations that are systematically varied according, for example, to the status of the addressee. Regardless of the type of task used, the grade 6 immersion students do not demonstrate a native-like command of this aspect of communicative competence. Very briefly, it may be noted that native speakers demonstrate a larger difference between formal and informal registers than do the immersion students (Allen et al. 1983). From a methodological perspective, it is clear that comparisons with native speakers are an essential aspect of the analysis of sociolinguistic competence, since it is often unclear whether any 'error' has actually been made.

(b) Accounting for the proficiency results

One way of viewing the proficiency findings in early French immersion is to accept them as the best possible outcomes that can be achieved in a classroom setting. This is, in fact, the reaction of Krashen (1984), for whom the goal of the language class is to produce "intermediates," whose acquisition of the second language is sufficiently developed that it can continue "on the outside" in interaction with native speakers. According to Krashen (1984:61):

> Second language acquisition theory provides a very clear explanation as to why immersion works. According to current theory, we acquire language in only one way: when we understand messages in that language, when we receive comprehensible input.

For Krashen, teaching that is focussed on the conveying of interesting and relevant messages will automatically provide the comprehensible input, or i + 1, a little beyond the learner's current stage of development, that promotes further second language acquisition (Krashen 1982). In his view, there is no need to practise since the ability to speak a language "develops on its own." Formal activities considered characteristic of traditional language classrooms, such as the memorizing of vocabulary, the study of grammar, and drills, are seen as making "a very small contribution to language competence in the adult and even less in the child" (Krashen 1984:61).

While we may agree with Krashen that native-like competence is a unrealistic goal for any language teaching situation where the students are isolated from contact with native speakers other than their teacher, there remains an important question as to whether we should be entirely satisfied with all the proficiency outcomes that have been documented in immersion. Although the mainly oral comprehensible input the students have received is clearly necessary for second language acquisition, the proficiency results suggest that it may not have been sufficient. We know, for example, that at age 6-7, grade 1 immersion students can accurately translate a French sentence containing a conditional form into English, but even after ten years of immersion in grade 10 we still find only a 56% accuracy rate in the oral production of conditionals (Harley and Swain 1984). One explanation for the potentially long-lasting discrepancy between global comprehension of language in context and accurate production lies in the characteristic redundancy of language. To comprehend the semantics of the messages they hear, immersion students may not need to process the language at a syntactic level (Swain 1985), and even in production the findings show that they may often get their message across without accurate use of semantically redundant grammatical features. Indeed teachers focussing on subject matter content may not even notice the errors that their students are making. Chaudron (1977) reported, for example, that late immersion teachers were more likely to correct content errors than language errors in subject matter classes. Students may also become adept at communication strategies that disguise a lack of grammatical competence. For example, the fact that in their oral French the students sometimes neutralize the distinction betwen le and la as may begin as a perceptual problem but eventually become a strategy to avoid making a choice in gender that escapes unnoticed by the teacher.

In contrast to Krashen, it may be hypothesized that within the confines of a classroom, where young students lack access to native speakers and the social motivation to speak like them, and where the range of speech acts that occur is likely to be limited in comparison to a natural language acquisition context, comprehensible input by itself is not enough. The mainly oral input received by students may not be as frequent or salient in specific grammatical or sociolinguistic domains of the second language as it could be, teachers may not be providing enough feedback or 'negative input' [4] (Schachter, 1983) aimed at alerting students to problems in their production, there may be little opportunity for students to express themselves freely in what Swain (1985) has called 'comprehensible output', and little need to produce certain forms in the speech acts that occur naturally in the course of the day. These considerations lead to the hypothesis that special efforts may be needed:

1. to provide more focussed second language input that promotes perception and comprehension of the formal and functional contrasts in French that continue to cause problems; and
2. to provide more opportunities for production of these different forms and functions in the realization of interesting, motivating tasks.

(c) Classroom observation to refine hypotheses

In preparation for an experimental study at the grade 6 level, designed to test the above hypotheses concerning the further instructional needs of immersion students, observations were conducted in immersion classrooms to determine the nature of the second language input already being provided and to assess the opportunities that students currently have to use French productively.

Nine grade 3 and ten grade 6 classes in three Ontario school boards were observed and tape-recorded for a full day each, and information about materials in use was obtained from each teacher. The recordings, which capture the public classroom talk of the teachers and students, have been transcribed, and 1 1/2 hours of the French portion of the day in each class has so far been analysed with respect to the following questions:

1. In view of proficiency findings indicating that the French verb system is particularly problematic for immersion students, what is the relative frequency of various verb tenses in immersion teacher talk and written French materials?
2. In what ways are teachers already focussing on problem areas of French grammar?
3. What kinds of opportunities do students have to talk in class?

4. Are there any changes in second language use from grade 3 to grade 6 that would indicate increased complexity in comprehensible input?

Frequency of verb forms in teacher talk

In a 1 1/2 hour period of class time, there is on average a relatively small proportion of use by the teachers of verb forms other than the present or imperative. Table 1 shows that past tenses , for example, represent only about 20% of all finite verbs used by the teachers in different subject areas (in the 1 1/2 hours all classes devoted at least some of the time to français, while some classes were also doing mathematics or social studies). A similar pattern of use may be observed for the grade 3 and the grade 6 teachers. When the use of past tenses is broken down into imparfait, passé composé, passé simple, and plus-que-parfait (see Table 2), it is evident that the passé composé is the most frequently used past tense in both grades (83% in grade 3 versus 64% in grade 6). However, along with an apparent decline in the proportion of passé composé used in grade 6, there is some increase in the use of the imparfait at this grade level. A closer analysis of the kinds of verbs used by the teachers in the imparfait shows that about two-thirds of them are non-dynamic verbs. Even in grade 6, therefore, the students appear to be getting relatively little exposure to the incomplete/complete aspectual distinction characteristically expressed by the use of the imparfait versus the passé composé in referring to actions.

An analysis currently in progress of verb tense usage by grade 6 teachers in two native French-speaking classes may reveal that the distribution of tenses in immersion teacher talk does not differ substantially from mother-tongue classrooms at this grade level. However, for immersion students, unlike native French-speaking students, virtually the only exposure to oral French is in the classroom, and the continuing relatively low frequency of past, future, and conditional tenses in grade 6 teacher talk goes some way towards explaining the problems that students have with the use of non-present tenses (see page 4).

Frequency of verb forms in written materials

Considerably more use of different tenses is evident in some of the texts [5] that the grade 6 students are using (grade 3 materials have not yet been analysed) and this may be an important source of new comprehensible input for the students. Table 3 shows that in their readers, grade 6 students are exposed to a much higher proportion of past tense use than in teacher talk. Relative to the amount of oral French that they hear from their teacher, however, it is likely that the average amount of text the students cover each day in their reader is much smaller. It is also evident (see Table 4) that the past tense usage to which they are exposed in their written materials differs from the oral input they receive from their teacher (further analysis of this phenomenon is being undertaken). It may be that in their readers the students are being exposed to so many different verb inflections at once that they tend to pass over them in extracting the essential meaning from other clues in the text. The use of more reading material could, however, be an important means of widening immersion students' exposure to the second language.

Activities focussing on French grammar

In eight of the grade 6 classes, but only three grade 3 classes, there was a period given to grammar activities. These grammar activities involved mainly the learning of formal paradigms and categories, and rules of written grammar. In the grade 6 classes, for example, there were two classes conjugating verbs, another class parsing sentences and identifying different types of object complement, a class labelling personal pronouns, and a class doing written grammar exercises involving verb agreement and the choice of auxiliaries ÊTRE and AVOIR in the passé composé. There was in general a much greater emphasis in these grammatical activities on the learning and categorising of forms than on relating the forms to meaning in context. However, an increased focus on formal grammar appears to be a distinct change from grade 3 to grade 6, designed to increase students' attention to grammatical features.

In other language arts activities, such as reading, students would in some classes in each grade be called upon to read aloud and teachers would on occasion correct pronunciation, misreadings, and grammatical errors (e.g. gender errors). The main emphasis of the teachers was on the comprehension of vocabulary in the stories being read, and this emphasis carried over into the comprehension exercises that often followed the reading.

Some language arts activities consisted of oral discussions and games that provided contexts for the use by students of specific verb forms. In general, however, such focussed opportunities for the use of specific grammatical forms were infrequent in either grade 3 or grade 6 classes, and such oral activities appeared to elicit little 'negative input' from teachers with respect to the students' language use.

In line with our hypotheses, these findings suggest that there may be room for more emphasis in the immersion context on providing language input that is explicitly designed to clarify the meaningful use of particular grammatical forms, and on fashioning communicative contexts in which students practise the productive use of such forms.

Opportunities for talk by students

An assessment of the talk turns that were taken by students in the various grade 3 and grade 6 classes (see Table 5) shows that there are relatively few occasions involving 'sustained' talk by students--i.e. talk of more than a single clause in length. In fact, a substantial proportion of the talk turns in both grades 3 and 6 consisted of 'minimal' one- or two-word responses to teacher initiations. In these instances, the teacher was typically seeking brief responses to questions of fact.

The analysis of talk turns did not reveal any significant differences between grade 3 and grade 6, suggesting that in absolute terms, the opportunities that grade 6 immersion students have for sustained talk in their second language have actually declined, since, in their case, the 1 1/2 hours analysed represents most of the French part of their day, as opposed to grade 3 for whom the 1 1/2 hours represents less than half of the time that they still spend immersed in French.

Refinement of hypotheses

The classroom observation data analysed so far (and further analyses are planned) do not disconfirm our initial hypotheses concerning the further language learning needs of the immersion students. Viewed from the theoretical perspective of Allen's (1983) three-level curriculum for second-language education (see Figure 1), the immersion classrooms we have observed may be described as focussing mainly on listening comprehension at the 'experiential', non-analytic level of the language curriculum, with more attention being paid to the 'structural-analytic' level in grade 6 than in grade 3, and with little attention apparently being given to the 'functional-analytic' level at either grade level. In Allen's view, all three curriculum components are necessary in any language teaching program, with varying emphases according to particular circumstances. Another way of formulating our hypotheses, therefore, would be to claim that immersion students would benefit from a greater emphasis on the functional-analytic component of the curriculum. At the same time, there appears to be room for different second-language-oriented activities at the structural-analytic level that are geared to specific needs of the immersion students with their English mother tongue background (e.g. to help them in mastering grammatical gender), and more room at the experiential level for sustained oral and written production of the second language by the students and for increased exposure to extended written text.

(d) Development of materials in accordance with hypotheses

A set of materials has been developed (Harley, Ullmann, and Mackay 1985) in an attempt to fill the functional-analytic gap with respect to the use of past tenses in immersion classroooms and to fill the hypothesized need for more sustained productive use of French by the students. Designed for flexible use in grade 6 early immersion, as well as in other types of French programs, the piloted materials provide input and practice aimed at teaching functional contrasts between the imparfait and the passé composé by means of interesting activities related to worthwhile subject matter themes. These oral and written activities include, for example, reading a simplified French-Canadian legend about a werewolf, discovering how the imparfait and the passé composé serve different functions in this legend, illustrating contrasting sentences such as _il dégringolait l'escalier_ / _il a dégringolé l'escalier_, fitting proverbs to the legend and to personal experiences, miming progressive uses of the imparfait in the context of a guessing game, creating new legends, producing individual albums of childhood memories, and reporting on 'research' concerning prehistoric animals. An explanatory handbook is provided for the teachers.

(e) Experimental testing of materials

Beginning in October, selected activities (including those mentioned above) will be used in a controlled eight-week experiment in six grade 6 immersion classes. The students' competence in using the imparfait and the passé composé will be compared in pre-tests and in immediate and delayed post-tests with that of six other immersion classes who are not exposed to the materials. Teachers using the materials will first be introduced to them in a full-day workshop and will be encouraged to continue teaching along the same lines. This experiment in one small area of French grammar will provide a test of our hypotheses with respect to the further needs of immersion students for functional-analytic language teaching in a specific domain.

Conclusion

In this paper, a research cycle has been presented which spans the range from analysis to diagnosis and proposed treatment. It is a cycle which we can foresee having to repeat many times in order to refine our hypotheses and accommodate learning needs in different age-groups and in different classroom contexts. For children, a major emphasis from the beginning on communicative use of the second language is clearly fundamentally sound. What is needed now is more classroom experimentation to test —in other program settings, too — the theoretical arguments that have been proposed to account for the proficiency findings in early total French immersion and that have provided the basis for the construction of materials designed to enhance the students' communicative competence in one specific area of grammar.

Footnotes

1 Much of the data on which this paper is based was collected in the context of the Development of Bilingual Proficiency Project (principal investigators J.P.B. Allen, J. Cummins, B. Harley, and M. Swain), funded by the Social Sciences and Humanities Research Council of Canada (Grant no. 431-79-0003). I am grateful to project staff, and especially Mary Lou King, Laurette Lévy and Françoise Pelletier, for their role in collecting, scoring, transcribing and analysing data.

2 Indirect evidence of the students' high level of comprehension is also provided by their performance on standardized tests of achievement (in English) in subjects they have been taught in French (Swain and Lapkin 1982).

3 For discussion of these and other problems in identifying and classifying errors, see Schachter and Celce-Murcia (1983).

4 Negative input is seen as including signals other than overt correction that the learner's utterance has been unsuccessful in some way (e.g. confirmation and clarification checks).

5 The analysis of these texts, regularly used in a number of the grade 6 classes, was based on a selection of pages at 7-page intervals in each text.

References

Allen, J.P.B. A three-level curriculum model for second-language education. **The Canadian Modern Language Review**, 1983, 40/1, 23-43.

Allen, J.P.B., Cummins, J., Mougeon, R. and Swain, M. The development of bilingual proficiency: Second year report. Toronto: Ontario Institute for Studies in Education, 1983.

Chaudron, C. Teachers' priorities in correcting learners' errors in French immersion classes. **Working Papers on Bilingualism**, 1977, 12, 21-44.

Commissioner of Official Languages. **Annual Report 1984.** Ottawa, 1984.

Harley, B. French gender 'rules' in the speech of English-dominant, French-dominant and monolingual French-speaking children. **Working Papers on Bilingualism**, 1979, 19, 129-156.

Harley, B. Transfer in the written compositions of French immersion students. In Dechert, H.W. and Raupach, M. (eds.) **Transfer in production.** Norwood, N.J.: Ablex (forthcoming).

Harley, B. and Swain, M. An analysis of the verb system used by young learners of French. **Interlanguage Studies Bulletin**, 1978, 3, 35-79.

Harley, B. and Swain, M. The interlanguage of immersion students and its implications for second language teaching. In Davies, A., Criper, C. and Howatt, A.P.R. (eds.) **Interlanguage.** Edinburgh: Edinburgh University Press, 1984.

Harley, B., Ullmann, R., and Mackay, L. Parlons du passé (Pilot version). Toronto: Ontario Institute for Studies in Education, 1985.

Krashen, S. **Principles and practice in second language acquisition.** New York: Pergamon, 1982.

Krashen, S. Immersion: Why it works and what it has taught us. In Stern, H.H. (ed.). 1984, 61-64.

Schachter, J. Nutritional needs of language learners. In M.A. Clarke and J. Handscombe (eds.) **On TESOL '82: Pacific perspectives on language learning and teaching.** Washington, D.C.: TESOL, 1983.

Schachter, J. and Celce-Murcia, M. Some reservations concerning error analysis. In B.W. Robinett and J. Schachter (eds.) **Second language learning: Contrastive analysis, error analysis, and related aspects.** Ann Arbor: University of Michigan Press, 1983.

Stern, H.H. (ed.) **The immersion phenomenon. Language and Society,** 1984, 12.

Swain, M. Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden (eds.) **Input in second language acquisition.** Rowley, Mass.: Newbury House, 1985.

Swain, M. and Lapkin, S. **Evaluating bilingual education: A Canadian case study.** Clevedon, U.K.: Multilingual Matters, 1982.

Figure 1

Allen's Three-Level Curriculum

(a) **Structural-analytic**

Focus on grammar and other formal features of language

Controlled grammatical teaching techniques

Medium-oriented practice

(b) **Functional-analytic**

Focus on discourse features of language

Controlled communicative teaching techniques

Medium and message-oriented practice

(c) **Non-analytic**

Focus on the natural unanalyzed use of language

Fully communicative, experiential teaching techniques

Message-oriented practice

From Allen (1983), p. 25

Table 1: **Average Use of Tenses by Grade 3 and 6 Teachers in Different Subject Areas**

	Subject	N. of classes	Finite Verbs	Use of Past Tenses %	Use of Future Tenses %	Use of Conditional Tenses %
			$\overline{X}$	$\overline{X}$	$\overline{X}$	$\overline{X}$
Grade 3	Français	9	325.4	12.7	7.7	1.5
	Maths	5	158.4	12.0	6.0	0.6
	Social Studies	2	74.5	19.8	16.6	0.0
Grade 6	Français	10	395.8	14.4	5.3	2.8
	Social Studies	5	186.2	21.4	6.8	4.0

Table 2: Breakdown of Past Tense Usage by Grade 3 and 6 Teachers*

	Total number of past tense forms used	Use of "imparfait" (% of past)	Use of "passé composé" (% of past)	Use of "passé simple" (% of past)	Use of "plus-que-parfait" (%of past)
	$\bar{x}$	$\bar{x}$	$\bar{x}$	$\bar{x}$	$\bar{x}$
GRADE 3 (8 classes)	57.5	17.3	81.5	0.5	0.7
GRADE 6 (9 classes)	76.7	29.4	64.5	3.4	2.6

*For 1 class at each grade level, the breakdown of past tenses was not available.

Table 3: **Tenses Used in Written Materials** (Grade 6)

Subject	No. of pages	Total N. of finite verbs	Present %	Past Tenses %	Future Tenses %	Condt'l Tenses %	Subj. & Participles %	Imperative %
Reader	14	372	29	51	3	3	11	3
Activities (Français)	6	124	40	27	5	-	14	14
Maths	18	291	46	8	1	1	7	38
Social Studies	15	364	34	32	7	2	11	12

Table 4: Breakdown of Past Tense Usage in Written Materials

SUBJECT	N. of pages	Total number of past tenses	Use of imparfait %	Use of passé composé %	Use of passé simple %	Use of "plus-que parfait"%
LECTURE 1	13	227	49	4	26	21
LECTURE 2	14	189	54	11	27	8
LECTURE 3	13	120	28	16	48	8
ACTIVITIES (FRANÇAIS)	6	33	58	33	9	-
MATHS	18	24	29	63	8	-
SOCIAL STUDIES	15	118	41	23	29	8

Table 5: Categories of Student Talk

	Total Number of utterances	Minimal (% of total)	Phrase (% of total)	Clause (% of total)	Sustained (% of total)	Sustained except read from text (% of total)
	$\bar{X}$	$\bar{X}$	$\bar{X}$	$\bar{X}$	$\bar{X}$	$\bar{X}$
Grade 3 (9 classes)	223.2	38.5	10.9	34.4	16.3	12.9
Grade 6 (10 classes)	226.7	39.5	11.7	30.1	18.7	14.9

Appendix B

School #126

Texts and Reference Materials

- Canac-Marquis, Jean 1979 Communication orale a in the series Le Français sous toutes ses formes, Centre Educatif et Culturel, Inc., Montreal.
- Canac-Marquis, Jean, Diane Séguin and Marielle Richler. 1980 Communication orale a : guide d'utilisation et texte des enregistrements, in the series Le Français sous toutes ses formes, Centre Educatif et Culturel, Inc., Montreal.
- Canac-Marquis, Jean, Jean Ethier and Denis Arpin. 1979 Communication orale a : entrainement à l'expression, in the series Le Français sous toutes ses formes, Centre Educatif et Culturel, Inc., Montreal.
- Canac-Marquis, Jean, Jean Ethier and Denis Arpin. 1979 Communication orale a : entrainement à la compréhension, in the series Le Français sous toutes ses formes, Centre Educatif et Culturel, Inc., Montreal.
- Carleton Board of Education. 1982 Immersion précoce français/arts 6e année, revised by Raymond Way, Carleton Borad of Education.
- Lédesart, R.P.L. and Margaret Lédesart (eds.) 1967 Harrap's New Shorter French Dictionary, revised edition, Harrap, London.
- Le Nouveau Bescherelle : L'art de conjuguer : dictionnaire de 1200 verbes. 1980 Editions Hurtubise H.M.H. Ltée, Montreal.
- Picard, M., G. Foen, E. Leroy and J. Dansereau. 1967-1973 La Grammaire à l'école active, 4th Volume, Centre Educatif et Culturel, Inc., Montreal.
- Rey, Alain, Josette Rey-Debove et al. 1971 Le Micro-Robert : dictionnaire du français primordial, Paris.
- **Readings**
- Canac-Marquis, Jean and Henriette Major. 1976 La Lecture sous toutes ses formes Level 2, Centre Educatif et Culturel, Inc., Montreal.
- Carrière, Jean. 1973 L'Enfant et la rivière, in the series Prenons la parole, Editions Ecole Active.
- Rennie, A. James. 1977 Un Joueur ambitieux, Maclean-Hunter Learning Materials, Toronto.
- Savoir Lire : sélections du Readers' Digest 1973 Translation and adaptation by IREC Inc., Reader's Digest.

- **Environmental Studies**

- Kiil, Susan, Diva Anderson et al. 1979 Ce Monde qui t'entoure : l'hémisphere occidentale, translation by Guy Vauban Ltée., McGraw-Hill Ryerson, Toronto.

- McConnell, G. Robert. 1975 Voici le Québec, Editions Aquila, Montreal.

- Also used were: Nouveau Larousse élémentaire Larousse, Paris. Unfortunately, full bibliographic references were not obtained at the time of the observations, and could not be found subsequently.

- **School #136**

- **Texts and Reference Materials**

- Anonymous group of teachers. 1977 J'explore le français Level 6, Editions Guérin, Montreal.

- Canac-Marquis, Jean. 1976 Code grammatical sélectif, Centre Educatif et Culturel, Inc., Montreal.

- Carleton Board of Education. 1982 Immérsion précoce français/arts 6e année, revised by Raymond Way, Carleton Board of Education.

- Girard, Denis (rev.), with the assistance of Gaston Dulong, Oliver van Oss and Charles Guinness. 1981 Cassel's French/English Dictionary, Cassell Ltd., London.

- Rey, Alain, Josette Rey-Debove et al. 1971 Le Micro-Robert : dictionnaire du français primordial, Paris.

- **Readings**

- Dubé, Gérard-A. and Andrée Soucie-Dubé. 1981 Contes d'ivoire, part of the series Méthode éclectique de lecture. Editions Guérin, Montreal.

- Dubé, Gérard-A. and Andrée Soucie-Dubé. 1980 Contes d'argent, part of the series Méthode éclectique de lecture, Guérin, Montreal.

- **Environmental Studies**

- Kiil, Susan, Diva Anderson et al. 1979 Ce monde qui t'entoure : l'hémisphere occidental. McGraw-Hill Ryerson, Toronto.

- Also used were Atlas Scolaire, Gage, and Dictionnaire Jeunesse, Centre Educatif et Culturel, Montreal. Unfortunately, full bibliographic references were not obtained at the time of the observations, and could not be obtained subsequently.

School #156

- **Texts and Reference Materials**

- Canac-Marquis, Jean and Henriette Major. 1977 La Lecture sous toutes ses formes Level 3, Centre Educatif et Culturel, Inc., Montreal.

- Fortin, Louis N. 1978 Pensées, proverbes et maximes Tome 1, Promotions Mondiales, Inc.

- Grévisse, Maurice. n.d. Précis de grammaire française, 27th edition Editions J. Duculot, Gembloux.

- Laurence, Jean-Marie with the collaboration of Aurèle Daoust. 1977 La Grammaire française fondamentale, Editions Guérin, Montreal.

- Le Nouveau Bescherelle L'Art de Conjuguer : dictionnaire de 1200 verbes. 1980 Editions Hurtubise H.M.H. Ltée., Montreal.

- Petit Larousse illustré 1977 Libraire Larousse, Paris.

- Picard, M., M. Cabau, B. Jughon and P. Gélinas 1967-1972 Le Langage à l'école active, 4th Volume, Centre Educatif et Culturel, Inc.

- Picard, M., G. Foen, E. Leroy and J. Dansereau. 1967-1973 La Grammaire à l'école active 4th Volume, Centre Educatif et Culturel, Inc.

- **Readings**

- Savoir Lire : sélections du Reader's Digest 1973 Translation and adaptation by IREC Inc., Reader's Digest.

- Also used were Cassell's Concise French/English Dictionary; Laurin, Jacques. Les verbes : la conjugaison rendue facile. Editions de l'homme; Le Nouveau Dictionnaire des Synonymes, published by Larousse; Le Nouveau Larousse Elémentaire; Camus, Albert Le Petit Prince; Forget, Solange. Dossier de l'élève - Le Petit Prince, Centre Franco-ontarien de Resources Pédagogiques; Phillips, June K. Petits contes sympathiques, National Textbook Co. Unfortunately, full bibliographic references were not obtained at the time of the observations and could not be obtained subsequently.

- **School #316**

- **Texts and Reference Materials**

- Canac-Marquis, Jean. 1979 Communication orale a, in the series Le Français sous toutes ses formes, Centre Educatif et Culturel Inc., Montreal.

- Clouthier, Berthe, Jacques Gosselin, André Mareuil and Josette Poulin. 1977 La Nouvelle grammaire de base, Editions Françaises.

- Genouvrier, Emile and Daniel Poulin. 1973 De la langue parlée à la langue écrite. Les Editions Françaises, Montreal.

- Giroux, Emile. 1970 Langue française Level 6, Editions FM.

- **Readings**

- Pouliot, Germaine. 1978 Le temps d'une rencontre Editions Projet, Montreal.

- Corriveau, Monique. 1979 Patrick et Sophie en fusée, Editions Héritage, Montreal.

- Darios, Louise. 1976 Reportages du chat Alexandre au Brésil, Editions Héritage, Montreal.

- Martel, Suzanne. 1979 Pi-Oui, Editions Héritage, Montreal.

- Anonymous group of teachers. Grammaire et orthographe Level 6B, Editions Guérin, Montreal; Saint-Jean, Albert Près de la fontaine, Editions Nelson, Toronto; Comtesse de Ségur Un bon petit diable; Lire et parler, Fernand Nathan, Paris. Unfortunately, full bibliographic references were not obtained during the observations and could not be obtained subsequently.

- **School #326**

- **Texts**

- Girard, Denis (rev), with Gaston Dulong, Oliver van Oss and Charles Guinness. 1981 Cassell's French/English Dictionary, Cassell Ltd., London.

- Genouvrier, Emile and Daniel Poulin. 1973 De la langue parlée à la langue écrite, with workbook, Les Editions Françaises, Montreal.

- **Readings**

- --------- 1974 Au pays des contes : récits et poèmes choisis, Editions Projets, Montreal

- Corriveau, Monique. 1969 Les jardiniers du hibou, Collection Karim, Editions Jeunesse Inc., Montreal.

- Dubé, Gérard-A and Andrée Soucie-Dubé. 1981 Contes d'ivoire, plus workbook, part of the series Méthode éclectique de lecture, Editions Guérin, Montreal.

- Manière, Renée. 1962 Conte du chien perdu, with workbook, Editions Education Nouvelle, Montreal.

- Roberge, Suzelle and Germaine Pouliot. 1985 Carrousel in the series Méthode dynamique du français Level 6, Editions Etudes Vivantes, Montreal.

- Also used were Larousse de poche français-anglais, Larousse, Paris; Lire et parler, Fernand Nathan, Paris; L'Invitation au voyage; A la découverte du monde. Unfortunately full bibliographic references were not obtained at the time of the observations and were not obtainable subsequently.

School #336

- **Texts and Reference Materials**
- Atkins, Beryl, Alain Duval et al. (eds.) 1978 Collins/Robert French/English English/French Dictionary, Collins, London.
- Bray, Alain de and Michel Therrien. 1981 Nouveaux exercises de grammaire française, with workbook, Editions Brault et Bouthillier, Ottawa.
- **Readings**
- Canac-Marquis, Jean and Henriette Major. 1977 La lecture sous toute ses formes Level 3, with workbook, Centre Educatif et Culturel Inc., Montreal.
- Pouliot, Germaine. 1977 Un ami sur la route, Levels 5 and 6, part of the series Méthode Dynamique, Editions Projets, Montreal.
- Also used were Compréhension de la langue française; Larousse de poche français-anglais Larousse, Paris; Aux Quatre vents. Unfortunately, full bibliographic references were not obtained at the time of the observations and could not be obtained subsequently.
- **School #346**
- **Texts**
- Bray, Alain de and Michel Therrien. 1981 Nouveaux exercices de grammaire française, with workbook, Edition Brault & Bouthillier, Ottawa.
- ?Faubert, Jacqueline. 1982-1984 Vive le français par coeur : grammaire et orthographe, Levels 6A and 6B, with workbook, Editions, Guérin, Montreal.
- Collin, Peter, Helen Knox, Margaret Ledésart and René Ledésart (eds.) 1982 Harrap's Shorter French/English Dictionary, Harrap, London.
- Rey, Alain and Josette Rey-Debove et al. 1971 Le Micro-Robert : dictionnaire du français primordial, Robert, Paris.
- **Readings**
- Pouliot, Germaine. 1978 Le temps d'une rencontre, Editions Projets, Montreal.
- Canac-Marquis, Jean. 1974 La Lecture sous toutes ses formes, Level 1 (this level is not certain), Centre Educatif et culturel, Montreal.
- Also used were Le Dictionnaire jeunesse, Centre Educatif et Culturel Inc., Montreal.Les Misérables;Le Cid; Evangeline; Contes de nöel; Monte Cristo; La Tulippe noire. Robinson Crusöe; L'Appel du grand-nord; Glozgab. Unfortunately, full bibliographic references were not obtained at the time of the observations and could not be obtained subsequently

- **School #356**

- **Texts and Reference Materials**

- Anonymous group of teachers. 1975-1979 J'Apprends à connaître la langue française Level 6, with workbook, Editions Guérin, Montreal.

- Girard, Denis (rev), with Gaston Dulong, Oliver van Oss and Charles Guinness. 1981 Cassell's French/English Dictionary, Cassell Ltd., London.

- Le Nouveau Bescherelle : L'art de conjuguer, dictionnaire des 1200 verbes. 1980 Editions Hurtubise H.M.H., Montreal.

- Pouliot, Germaine. 1978 Le Temps d'une rencontre, Editions Projets, Montreal.

- Also used were L'Orthographe d'instinct, Editions Auto-correct-art, Quebec. Larousse des débutants; Larousse français/anglais, Larousse, Paris. Unfortunately, full bibliographic references were not obtained during the observations and could not be obtained subsequently.

Appendix C

ASPECTS OF THE SOCIOLINGUISTIC PERFORMANCE OF EARLY AND LATE FRENCH IMMERSION STUDENTS

Merrill Swain   Sharon Lapkin
The Ontario Institute for Studies in Education
Toronto, Ontario

Over 125,000 English-speaking students in Canada are now learning French as a second language through 'immersion'. That is, for a period of time which may extend over a number of years, these students are being taught a major portion of their curricular content using French as the language of instruction. There have been a considerable number of studies undertaken which have examined the first language proficiency of these immersion students, as well as their academic achievement and cognitive growth. Even though much of their education has been in a second language, the results of these studies have shown that the immersion students suffer no long-term detriment in any area of scholastic achievement or in their cognitive development when compared to students educated entirely in English (see Swain and Lapkin, 1982 for a review). This paper, therefore, is not concerned with the academic side of the immersion students' performance. Rather, it will consider some aspects of their second language development.

Specifically, this paper will examine aspects of the spoken and written sociolinguistic performance of grade ten immersion students. Sociolinguistic performance reflects the degree to which specific utterances are appropriate given the topic, the status of the participants, the purpose of the interaction and other aspects of the sociolinguistic context. Interest in the acquisition of sociolinguistic skills has been spurred by the current focus in second language pedagogy on communicative language teaching. Communicative language teaching, it has been argued (Canale and Swain, 1980; Canale, 1983), must address itself minimally to teaching grammatical, discourse, sociolinguistic and strategic competence. The results of communicative language teaching must, therefore, be examined in light of its impact on each of these four theoretically posited dimensions.

Previous research concerned with assessing the second language proficiency of the immersion students has focussed mainly on grammatical and discursive aspects of listening and reading comprehension tasks (Green and Lapkin, 1984). However, given that one of the goals of French immersion education in Canada is to provide participating students with the ability to communicate with French-Canadians, then being able to use French in sociolinguistically appropriate ways is an important aspect of what the students need to acquire.

The students involved in the current study were enrolled in an immersion program that began either when the students were five years old (early immersion) or when they were twelve years old (late immersion). Our expectation is that the students who began their French immersion program at a younger age will perform more like native speakers of the target language than will those who started at an older age. This prediction is based not only on the fact that the younger students will have had more second language input generally, but on the assumption that they will have had a wider range of in-class sociolinguistic experiences.[2] The way in which the immersion teachers use the target language in class will, therefore, be expected to affect the students' sociolinguistic performance.[3]

Also expected to play a role in the sociolinguistic performance of the immersion students is language transfer.[4] It is interesting to note the relative paucity of research concerned with sociolinguistic transfer: in a recent volume edited by Gass and Selinker (1983), only one article (Olshtain, 1983) is relevant to this topic.

Students tested: The students who were tested in this study were involved in one of two early immersion programs or a late immersion program. Students in the late immersion program and in one of the early immersion programs involved in the study were enrolled in the Carleton Board of Education, a board located in the capital of Canada, Ottawa. Students in the other early immersion program involved in the study were enrolled in the Toronto Board of Education. Also tested for comparison purposes was a group of bilingual francophone students from Ottawa.[5]

The early immersion students had had all of their instruction in French during the first three years of schooling, followed each year by decreasing amounts of time when French was used as the medium of instruction. By the time of testing (April, 1983) when the students were in grade ten, approximately one half of their instructional time was in French.

The late immersion students began their intensive exposure to French in grade seven. In grades seven and eight, approximately 80% of their instructional time was in French, which was reduced to about 50% in grades nine and ten. Prior to beginning their immersion program, these students had studied French as a subject for 20 minutes a day (kindergarten to grade five) to 40 minutes a day (grade six).

The selection of students for testing was made after certain students had been excluded from each class list (three early immersion classes, two late immersion classes and two classes from a French language school). In the immersion classes, students were excluded if they had not been enrolled in their respective program from its beginning. Thus, for example, late immersion students who had had some early immersion experience were excluded. Also, immersion students with French in their home background were excluded. From the French language school classes, students who indicated that the language spoken at home was not French or who had not been enrolled in a French language school throughout their schooling were excluded.

Once the exclusions were made, a stratified random sample of students was selected. Students were selected randomly from within three categories: those obtaining high scores, those obtaining scores in the middle range and those obtaining low scores on a French cloze test given in grade nine. 25% of the sample were chosen from the first and third groups and 50% of the sample were chosen from the second group. The final sample consisted of thirteen early immersion students from Ottawa, seven early immersion students from Toronto, eleven late immersion students, and six francophone students.

Tests: The students tested were given an extensive battery of tests of French proficiency (see Lapkin and Swain (1984a; b) for a complete description of the tests used), two of which are relevant in the present context. The first test, À vous la parole, is a test of communicative performance (Swain, 1984) focussing particularly on the assessment of speaking and writing skills. The second test is a speaking test intended to measure certain aspects of sociolinguistic performance.

The first test, À vous la parole, is a theme-oriented test. The theme of À vous la parole relates to two fictitious summer employment projects for Canadian youths fifteen years and older. The title was taken from a letter by a young Montrealer included in the student booklet in which he claims that youths face discrimination in the job market and suggests that it is time for adults to "let us have our say" in decision-making that affects youth. The two projects relate to different interest areas. The first involves organizing a series of rock concerts in Sudbury, Ontario; the second is located in the historic site of Louisbourg, Nova Scotia, where participating students assist with farm and animal care. A series of writing and speaking tasks based on the projects are given to the students being tested. The writing tasks include a letter, a composition, an informal note addressed to peers and a technical exercise (involving transforming point form information into expository text). The oral tasks include a group discussion among four students and a simulated job interview. In this paper only the results associated with sociolinguistic performance will be considered. The tasks which included scores for sociolinguistic performance were the note-writing task and the job interview.[6]

In the note-writing task, two aspects of sociolinguistic performance were assessed. The first was the appropriateness of the opening phrase. The students had all been informed that the note was to be informal and was to be left on the cafeteria bulletin board. It was therefore appropriate to begin the note with some sort of attention-getting device such as Salut toi!, POUR TOI ADOLESCENT, or ATTENTION!. Such openings were used by the francophone students in their notes. One point was awarded when an appropriate opening was used.

The second aspect of sociolinguistic performance assessed in the note was the use of the conditional. Effective notes written by the francophone students indicate that they can be written with or without the conditional verb form. However, a number of the notes written by the francophone students did include the use of the conditional. The conditional is used for two purposes in French: (1) for grammatical reasons to express the hypothetical mode, often with a co-occurring imperfect tense (e.g. Si tu étais employé, tu DEVRAIS t'occuper de l'éclairage); (2) for sociolinguistic reasons, to attenuate a request or command (e.g. POURRAIS-tu offrir une place dans ton auto?). In the latter case, the conditional is grammatically optional because the modal auxiliary alone could suffice (PEUX-tu...), although it does not attenuate the request as much as the conditional. An analysis of the use of the conditional in the notes written by the immersion students relative to native speakers was carried out.

In the job interview, one aspect of sociolinguistic performance was analyzed descriptively: the use of 'generic' personal subject pronouns.[7] Generic pronouns refer to a group, or to people in general, rather than to specific individuals. The generic pronouns in English are you and one. (E.g. To get to the library, you turn right at the next corner.) In French, this function can be expressed by nous, vous, on and tu depending on the degree of formality required by the sociolinguistic context. The analysis of students' use of generic personal subject pronouns in French is based on the entire job interview.

The second test -- the sociolinguistic speaking test -- consists of presenting a series of twelve situations using slides and audio accompaniment describing the situation. Each situation involved either making a request, a complaint, or a suggestion at one of two levels of formality. The test begins with the tester explaining to the student being tested how different registers of speech may be used in different situations and illustrating this with an example. The student then watches a set of two slides and listens to the synchronized description. With the showing of each slide, the student responds in the most appropriate way as if addressing the person shown in the slide. For example, in one set of slides, the first slide shows two children in the school library who are slightly younger than the student being tested. The student hears a description in French that says "You're in the library to study. But there are two persons at the next table who are speaking loudly and are bothering you. You decide to ask them to make less noise. What would you say if the two persons were friends of yours?". To change the level of formality, the second slide shows two adults in the library, and the final question is "What would you say if the two persons were adults that you don't know?".

There were two objectives of the scoring. One objective was to determine how the students performed in only the formal situations. The second objective was to determine the extent to which students could vary their language use appropriately in response to the social demands of the different situations. In other words, the scores were to indicate the student's ability to use linguistic markers of formal register in formal situations and to refrain from using them in informal situations. Thus, for each situation, a student's response was scored for the presence (=1) or absence (=0) of seven markers of formal register. The seven formal features were: (1) the use of an introductory politeness marker such as pardon or madame in the utterance opening; (2) the use of vous as a form of address; (3) the use of question forms with est-ce que or inversion; (4) the use of the conditional verb form; (5) the inclusion of an attenuating phrase (e.g. un petit peu); (6) the use of politeness formula such as s'il vous plaît; and (7) the use of sophisticated vocabulary[8] (e.g. l'obligeance).[9] Two scores are reported based on these analyses: an 'absolute' score and a 'contrast' score. The absolute score is the average score obtained on each feature in the formal situations only. The contrast score is obtained by subtracting the average number of, for example, uses of vous used across the informal situations from the average number of uses of vous used across the formal situations. This contrast score reflects the ability to differentiate between the two types of situations, a higher score reflecting a greater tendency to differentiate linguistically between the formal and informal setting than does a lower score.

Results: The results from the note-writing task are presented first, followed by the results from the two speaking tasks: the job interview and the sociolinguistic speaking test. In examining the results it must be remembered that what is being considered here is sociolinguistic behaviour. Such a concept cannot be considered in an absolute sense, but must be considered relative to native speakers of the language. That is to say, when it comes to judging sociolinguistic performance, second language learners' performance has to be based on a comparison of what native speakers of the same age do with the language in the same situation.

1. Note-writing task: (i) The first aspect of the note-writing task examined was the use of an appropriate opening, one that sought to attract the reader's attention by being both short and informal. There is no reason to expect that the use of an attention-getting device in an informal note would be any different when writing in French than when writing in English. That being the case, one would predict that no differences between

the groups would exist: the immersion groups simply apply their sociolinguistic knowledge in another language context. As shown in Table 1, the results support this prediction, there being no statistically significant differences among the francophone ($\bar{x}$ = .50), Toronto early immersion ($\bar{x}$ = .57), Carleton early immersion ($\bar{x}$ = .69) and Carleton late immersion ($\bar{x}$ = .64) groups.

(ii) The second aspect of sociolinguistic performance examined in the notes the students wrote was the use they made of the conditional for either the hypothetical (grammatical) or attentuating (sociolinguistic) functions. Table 2 shows the number of students from each group who used at least one conditional in their note.[10] One-half of the francophone students and approximately one-third of the early immersion students used a conditional in writing their note, while only approximately one-fifth of the late immersion students did so. These data suggest that the early immersion group is more similar to the francophone comparison group than is the late immersion group.

However, when the purposes for which the conditional is used are considered, the two immersion groups' performance appears similar, and substantially different from that of the francophone group. On the one hand, as Table 2 shows, each conditional a francophone student used serves the function of attenuating — making more polite — what is being said. For example, one francophone student wrote, J'écris cette note pour essayer de trouver une personne qui POURRAIT me conduire... Another wrote, Tu sais, j'AURAIS besoin de ton aide... .

On the other hand, both early and late immersion students use the conditional with its hypothetical meaning (for example, one early immersion student wrote, Ce SERAIT favorable si tu savais comment utiliser...) Additionally they are more likely to use hypothetical conditionals than attenuating conditionals. These findings complement those of Harley and Swain (in press) who concluded from their examination of the use of the conditional among early immersion students at several grade levels that "...the students are much more likely to produce conditionals in an obligatory context than they are to produce them in an optional context."

2. Job interview: (i) The use of selected personal subject pronouns used generically or non-generically in the job interview were tabulated for each program group: tu, on+ which refers to on when it is used for the non-generic nous,[11] on- which refers to on when it is used generically (e.g. Pour aller en Europe, ON traverse l'Atlantique.), nous and vous. The results appear in Table 3 where occurrences of each pronoun have been divided by the total number of such subject pronouns used by each program group. Thus, for example, tu accounts for 31% of all the generic personal subject pronouns used by early immersion students.[12]

The results shown in Table 3 suggest that early immersion students are more similar to native speakers than are late immersion students. This is so in three obvious ways. First, neither the francophone nor early immersion students used vous generically, while late immersion students did. Secondly, on used generically was almost never (2%) used by francophone students, was sometimes (26%) used by early immersion students, and was as likely as not (48%) to be used by late immersion students. Thirdly, francophone and early immersion students used tu generically considerably more frequently (63% and 31% respectively) than late immersion students (8%). Thus, although early immersion students differ somewhat from native speakers in their use of generic pronouns, their pattern of usage more closely resembles that of native speakers than does that of late immersion students.

3. Sociolinguistic speaking test: (i) The first aspect of the sociolinguistic speaking test examined was the use of an introductory politeness marker such as pardon or madame. Such markers are used in both languages, and are basically dependent on knowing the appropriate vocabulary. As Table 4 indicates, there are no significant differences in the extent to which immersion and francophone students make use of an

introductory politeness marker in formal situations. They use such markers in formal situations from two-thirds (francophone students) to three-quarters (immersion students) of the time. Furthermore, as Table 5 indicates, there are no significant differences in the extent to which they differentiate between formal and informal situations, all three groups making major adjustments. For example, in the formal version of one situation, an early immersion student said, PARDONNEZ-MOI, mais est-ce que tu peux chuchoter parce que je veux étudier?; while in the informal version, the same student said, Est-ce que tu peux chuchoter? Je veux étudier.

(ii) The second aspect of sociolinguistic performance examined was the appropriate use of vous and tu. Judging from native speaker behaviour, the formal situation called for the use of vous (see Table 4). Furthermore, as seen in Table 5, the large contrast score of the francophone students indicates that the use of tu was called for in the informal setting. As can be seen in Table 4, however, early immersion students use vous only about one-half the time in the formal setting, creating significant differences between themselves and both late immersion and francophone students. No significant difference exists between late immersion and francophone students on this measure. In this case, then, the inappropriately greater use of vous seen in the job interview by the late immersion students has worked in their favour in this task by bringing them closer to the usage pattern of native speakers, although they still do not differentiate between formal and informal settings to the same extent as francophone students do (see Table 5). In effect, the information combined from Table 4 and 5 suggests that late immersion students overuse vous in the informal situation, while early immersion students overuse tu in the formal situation relative to native speakers.

When the vous/tu results from the two tasks -- the job interview and the sociolinguistic speaking test -- are considered together, they appear contradictory in that in the job interview the language use of the early immersion students approximates more closely that of the francophone students than does that of late immersion students while on the sociolinguistic speaking test the reverse is true. In effect, it appears that early immersion students are 'doing as their teachers do, not as their teachers tell them to', while the late immersion students are 'doing as their teachers tell them to, not as their teachers do'. That is to say, based on our informal classroom observations, it would appear that teachers tend to address individual students as tu. The use of vous in the classroom setting appears to be largely reserved for addressing groups of students, thus signalling its use as a plural form. There are relatively few opportunities in the classroom for students to observe the use of vous as a politeness marker used in differential status situations, although as the students get older teachers may insist that students use vous when addressing them. At the time of testing, early immersion students have been exposed to these patterns of usage for eleven years; late immersion students, for only four years. Additionally, both groups of students have in all likelihood been taught a rule that goes something like, "vous is used with adults, and tu is used with friends and animals". Such a rule is obviously too simple to cover the uses of vous and tu involved in the job interview. It would, however, be of considerable help in performing appropriately on the sociolinguistic speaking test. The impact of teaching sociolinguistic rules, it is being suggested, is greater on late immersion students than on early immersion students in part because of the much more restricted exposure that the late immersion students have had with the language. In addition, late immersion students are more analytical relative to early immersion students at the time when their intensive exposure to French begins. The fact that in the more subtle aspects of the use of tu and vous where a rule might be difficult to formulate[13] as, for example, when used as generic pronouns, the early immersion students are more native-like in their performance reinforces this interpretation.

(iii) The third aspect of sociolinguistic performance examined was the use of question forms with est-ce que or inversion. Putting requests, suggestions or complaints in the form of a question appears to be a strategy equally as appropriate in English and French. As Table 4 shows, it is a frequently used strategy, and one that is used with equal frequency by all three groups of students in a formal situation. Furthermore, the lower contrast scores shown in Table 5 relative to those of the previous two measures indicate that although there is some differentiation between the formal and informal settings, there is still considerable use made of questions in the latter. For example, one late immersion student said when speaking to an adult, EST-CE QUE ça serait possible que vous m'aidez avec les mathématiques que j'ai manquées?, and when speaking to his peer, PEUX-TU m'aider avec les mathématiques que j'ai manquées?. In another setting however, the same student said to an adult POUVEZ-VOUS arrêter de me pousser?, but said to his peer ARRÊTE de me pousser. As Table 5 indicates, the groups do not differ significantly in the degree to which they differ in their use of est-ce que and inversion in the formal and informal settings.

(iv) The fourth aspect of sociolinguistic performance examined in this speaking test was the use of the conditional. It will be recalled that in the note-writing task, the early immersion students made more use of the conditional than did the late immersion students, and the francophone students made more use of the conditional than did either the early or late immersion students. Although the same trend is noted in Table 4 where the use of the conditional in the formal situations is shown, the difference between early and late immersion groups is not significant. Both groups, however, make significantly fewer uses of the conditional than the francophone group. Furthermore, as shown in Table 5, the two immersion groups do not vary their use of the conditional much between the formal and informal setting, and do so significantly less than do the francophone students.

(v) The fifth aspect of sociolinguistic performance examined in the test was the inclusion of an attenuating phrase such as UN PEU MOINS de bruit (late immersion student). According to Table 4, relatively little use of such attenuating phrases was made by any of the three groups in the formal situations, and as Table 5 indicates, relatively little change occurred in the use of this feature in the informal settings. No significant differences among the groups on the formal items alone, nor on the differential use of attenuating phrases in formal and informal settings were observed.

(vi) The sixth aspect of sociolinguistic performance examined was the use of politeness formula such as s'il te plaît and s'il vous plaît. Using s'il te plaît or s'il vous plaît is in principle no different than using 'please' in English, suggesting that there would not be differences between the groups on this feature. As Table 4 indicates, there were no significant differences noted between the groups in the frequency with which they made use of a polite formula in formal situations. While early immersion and francophone students tended to reduce to a similar extent the degree to which they used polite formula in informal situations, late immersion students did so to a significantly lesser degree than francophone students.

(vii) The seventh and final aspect of sociolinguistic performance which was examined in the speaking test was the use of sophisticated vocabulary. As with the use of attenuating phrases (see (vi) above), little use was made of sophisticated vocabulary. However, from Table 4 it can be seen that late immersion students used significantly fewer sophisticated vocabulary items than early immersion and francophone students, and that early immersion students used significantly fewer than francophone students. Table 5 indicates that virtually all the sophisticated vocabulary used was used in the formal setting, such that the contrast scores reflect the same significant differences as in Table 4.

Conclusions and Discussion: There are two general conclusions that can be drawn from the results. The first conclusion is that, with the exception of the use of vous and tu in the sociolinguistic speaking test, wherever there are significant differences noted between at least two of the groups on a particular sociolinguistic feature, the trend is always for early immersion students' sociolinguistic performance to be more like that of francophones than is that of late immersion students. Sometimes differences between the two immersion groups do not reach statistical significance, but the trend of early immersion students being more like francophone students than late immersion students are, is nevertheless evident.

The results, therefore, correspond in general to our expectation that the early immersion students' sociolinguistic behaviour would more closely approximate that of francophone students than would that of late immersion students. This expectation is based on the assumption that the longer period of exposure to the second language enjoyed by the early immersion students would provide more opportunities for input in a greater variety of situations.

When we examine those aspects of sociolinguistic performance where the trend is for early immersion students to be more like francophone students than are late immersion students, as compared to those aspects of sociolinguistic performance where there are no differences between the groups of students, two explanations are suggested: linguistic complexity and language transfer.

In particular, the trend is noted among the more subtle, non-categorical rule-based, non-formulaic features of sociolinguistic performance that we measured, as in, for example, the appropriate use of generic pronouns and conditionals. With more formulaic features such as the use of an appropriate opening, est-ce que, an attenuating phrase or s'il vous plaît, no differences among the groups were found. Thus, linguistic complexity would appear to play a role. So also would the first language.

We have indicated throughout this paper that the functions served by the aspects of sociolinguistic performance we have measured are functions that are present in both English and French contexts. They are, however, differently manifested in French and English.[14] The difference may be as minor as the translation of an isolated lexical item —from 'sir' to 'monsieur', from 'please' to 's'il te plaît', from 'a little bit' to 'un petit peu'; or as major as producing an appropriate complex morpho-syntactic structure in the target language -- from 'Would you explain...' to, for example, 'Pourriez-vous expliquer...' or 'Est-ce que tu pourrais expliquer...'. Immersion and francophone students tend to perform similarly when the difference between English and French is limited to the use of an appropriate word or phrase. These are important aspects of sociolinguistic performance, and the apparent ease with which they are applied in a second language context has largely been ignored in current research on language transfer.

The second conclusion to be drawn from the results is that teachers have an important role to play in developing the sociolinguistic competence of their students. This is particularly well demonstrated by the way in which the different groups use vous and tu. On the one hand, as discussed in the results section, the early immersion students appear to be 'doing as their teachers do, not as their teachers tell them to'. They appear to have learned by induction from the teachers' language the appropriate use of vous and tu. The problem is that the classroom environment is limited in the extent to which a distinction is made between vous and tu on the basis of differential status, and hence, what the early immersion students have learned is inappropriate to the wider out-of-school context. The late immersion students, on the other hand, appear to be 'doing as their teachers tell them to, not as their teachers do'. They appear to have learned by deduction from the rules the teachers have given them. The problem is that the rules cannot cover the complexity of appropriate usage patterns.

The solution, perhaps obvious by now, is that teachers are needed to provide both sociolinguistic rules and examples of language used appropriately in a variety of contexts.[15] Our prediction is, however, that the relative importance of one over the other will interact in important ways with the out-of-class language environment of the students and the age of the learner.

FOOTNOTES

1. We would like to express our thanks to Laurette Lévy and Diane Boucher for their help in the data analyses undertaken, and for their constant native speaker input to our inevitable non-native intuitions. We would also like to thank Susanne Carroll, Birgit Harley and Raymond Mougeon for commenting on an earlier draft of this paper, and Ian Gertsbain for his insights based on his experience in teaching sociolinguistic competence.

2. The use that French immersion students make of French outside of class is quite limited (see, for example, Lapkin et al, 1983).

3. Little research has been undertaken on immersion teacher talk. A study on this topic is currently underway in the OISE Modern Language Centre.

4. By language transfer we mean the use in some way of native language knowledge in second language performance.

5. For purposes of this study, francophone students are those for whom French is the language spoken at home and who have attended French language schools throughout their schooling.

6. Each task was initially scored for at least two aspects of communicative competence (discourse, sociolinguistic, grammatical or strategic).

7. The other personal subject pronouns (je, il, ils, elle and elles) and the non-generic uses of tu, nous and vous were also examined. As no obvious differences in their pattern of use among students in the three programs were noted, these data are not discussed further.

8. Sophisticated vocabulary is defined as the ability to deal with relatively infrequent vocabulary items. The term was operationalized by awarding one point to vocabulary not appearing in the Français fondamental list (Gougenheim et al., 1964).

9. In the following sentence, all such markers are present: Madame, auriez-vous l'obligeance de faire un petit peu moins de bruit, s'il vous plaît.

10. The data from the two early immersion programs have been combined in order to increase the sample size.

11. The criteria that were used to determine whether on was being used non-generically were essentially contextual ones, for example, presence of the possessive notre or nos, direct reference to self and other individuals.

12 The data from the two early immersion programs have been combined in this and subsequent analyses because, as indicated in the results from the note-writing task (appropriateness of opening), no differences between the two groups were noted.

13. Or indeed, might not even be known.

14. They may also be differentially distributed in subtle ways. Thus, for example, although 'sir' and *'monsieur'* and 'madam' and *"madame"* are translation equivalents, as polite appellatives 'madam' may not have the same distribution in English as *'madame'* has in French; whereas 'sir' and *'monsieur'* may be more similarly distributed. If these differences exist and have an impact on second language sociolinguistic performance, our measures were not fine-grained enough to detect them.

15. Raymond Mougeon (personal communication) comments: "Providing increased opportunities to use language in sociolinguistically meaningful situations and making sociolinguistic rules explicit will certainly help the late immersion students, but the fact that overall these students are exposed to and use the French language less often than early immersion students may prove to be an insurmountable obstacle with respect to aspects of sociolinguistic competence which are dependent on the mastery of more subtle, complex aspects of the French language. In Mougeon, Heller, Beniak and Canale (in press), we observed that the English dominant students of Ontario's French language schools (students whose use of French is mostly restricted to the school setting) are not sociolinguistically and linguistically as proficient as their French dominant counterparts (students who not only use French at school, but in the home as well). Our English dominant students are in a sense like the late immersion students in that they are confronted with the dual problem of not having had enough opportunities overall to use and hear French (to ensure mastery of more difficult elements of French) and of not having had enough time to use and hear French in different sociolinguistic settings (to ensure adequate development of sociolinguistic competence)."

REFERENCES

Canale, M. From communicative competence to communicative language pedagogy. In J. Richards and R. Schmidt (eds.) Language and Communication. London: Longman Group Ltd., 1983.

Canale, M. and Swain, M. Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1980, 1, 1-47.

Gass, S. and Selinker, L. (eds.), Language Transfer in Language Learning. Rowley, Mass: Newbury House, 1983.

Gougenheim, G., Rivenc, P., Michéa, R. and Sauvageot, A. L'Élaboration du français fondamental. Paris: Didier, 1964.

Green, D.Z. and Lapkin, S. Communicative language test development. In P. Allen and M. Swain (eds.), Language Issues and Education Policies, ELT Document 119, Oxford: Pergamon Press, 1984, 129-147.

Harley, B. and Swain, M. The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper and A.P.R. Howatt (Eds.), Interlanguage, Edinburgh University Press, in press.

Lapkin, S. and Swain, M. Second language maintenance at the secondary school level: Final report to the Carleton Board of Education. Toronto: Ontario Institute for Studies in Education, 1984a.

Lapkin, S. and Swain, M. Second language maintenance at the secondary school level: Final report to the Toronto Board of Education. Toronto: Ontario Institute for Studies in Education, 1984b.

Lapkin, S., Swain, M., Kamin, J., and Hanna, G. Late immersion in perspective: The Peel study. Canadian Modern Language Review, 1983, 39, 182-206.

Mougeon, R. Heller, M., Beniak, E. and Canale, M. Acquisition et enseignement du français en situation minoritaire: le cas des franco-ontariens. Canadian Modern Language Review, in press.

Olshtain, E. Sociocultural competence and language transfer: The case of apology. In S. Gass and L. Selinker (eds.) Language Transfer in Language Learning. Rowley, Mass: Newbury House, 1983.

Swain, M. Large-scale communicative language testing: A case study. In S. Savignon and M. Berns (eds.), Initiatives in Communicative Language Teaching: A book of Readings. Reading, Mass: Addison-Wesley, 1984, 185-201.

Swain, M. and Lapkin, S. Evaluating Bilingual Education: A Canadian Case Study. Clevedon, England: Multilingual Matters, 1982.

TABLE 1

MEAN SCORES ON THE "APPROPRIATE OPENING" MEASURE OF THE NOTE-WRITING TASK (λ VOUS LA PAROLE)

	Late Immersion N=11		Early Immersion (Carleton) N=13		Early Immersion (Toronto) N=7		Francophones N=6		Significant Difference
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	
Appropriate Opening (Max=1)	.64	.50	.69	.48	.57	.54	.40	.55	NS

TABLE 2

NUMBER OF STUDENTS USING CONDITIONAL FORMS IN THE NOTE - WRITING TASK (À VOUS LA PAROLE)

	Late Immersion N=11	Early Immersion (Toronto and Carleton) N=20	Francophone N=6
No. using conditionals	2	7	3
No. using hypothetical conditionals	2	6	0
No. using attenuating conditionals	1	2	3

TABLE 3

PERCENTAGE OF SELECTED PERSONAL SUBJECT PRONOUNS USED GENERICALLY (TU, ON-, NOUS, VOUS) AND NON-GENERICALLY (ON+) IN THE JOB INTERVIEW

	Late Immersion N=11	Early Immersion (Toronto and Carleton) N=20	Francophone N=6
TU	8	31	63
ON+	28	39	35
ON-	48	26	2
NOUS	8	4	0
VOUS	8	0	0

TABLE 4

SOCIOLINGUISTIC SPEAKING SCORES IN FORMAL SITUATIONS ON THE SOCIOLINGUISTIC SPEAKING TEST

	Late Immersion (L) N=11		Early (Immersion (E) (Toronto and Carleton) N=20		Francophone (F) N=6		Significant Differences
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	
Introduction	.77	.32	.78	.28	.64	.25	NS
Vous	.75	.34	.51	.34	.90	.11	E<L*; E<F*
Interrogative	.86	.13	.88	.13	.91	.14	NS
Conditional	.05	.15	.13	.28	.40	.23	L<F***; E<F*
Attenuating Phrase	.06	.11	.08	.10	.12	.14	NS
Polite Formula	.26	.23	.28	.23	.20	.13	NS
Sophisticated Vocabulary	.02	.05	.08	.13	.18	.11	L<E*; L<F***; E<F*

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

TABLE 5

SOCIOLINGUISTIC SPEAKING CONTRAST SCORES ON THE SOCIOLINGUISTIC SPEAKING TEST

	Late Immersion (L) N=11		Early Immersion (E) (Toronto and Carleton) N=20		Francophone (F) N=6		Significant Differences
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	
Introduction	.61	.25	.61	.24	.45	.25	NS
Vous	.58	.31	.37	.32	.89	.12	L<F**; E<L*; E<F**
Interrogative	.33	.15	.25	.20	.30	.18	NS
Conditional	.05	.15	.08	.16	.28	.25	L<F*; E<F*
Attenuating Phrase	.06	.11	.03	.06	.09	.10	NS
Polite Formula	.08	.12	.18	.20	.17	.11	L<B*
Sophisticated Vocabulary	0	0	.07	.12	.15	.08	L<E*; L<F**; E<F*

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

APPENDIX D

THE CASE FOR FOCUSSED INPUT: CONTRIVED BUT AUTHENTIC -- OR, HOW CONTENT TEACHING NEEDS TO BE MANIPULATED AND COMPLEMENTED TO MAXIMIZE SECOND LANGUAGE LEARNING

Merrill Swain

The Ontario Institute for Studies in Education

252 Bloor Street W.,

Toronto, M5S 1V6.

DRAFT: COMMENTS APPRECIATED

Plenary paper presented at TESL Canada '87 Conference, Vancouver, B.C., March 1987

Prologue

Sometime last November, Nick Collins phoned me and asked for a title for this talk. As I wasn't exactly sure what I was going to talk about, coming up with a title was somewhat problematic! I knew I wanted to talk about the fact that 'comprehensible input', a la Krashen, wasn't enough to account for second language acquisition. I have data that show that it isn't enough. So I thought I'd provide a title that implied that -- which is the one that appears in the program: " The Case for Focussed Input: Contrived but Authentic". I figured I could work out the details later.

Now that I've had more time to consider what I want to say, I feel the need to add a subtitle to that title which specifies more precisely my topic. The subtitle I'd like to add to "The Case for Focussed Input: Contrived but Authentic" is "How Content Teaching Needs to be Manipulated and Complemented to Maximize Second Language Learning".

Introduction

The purpose of this paper is to suggest that not all content teaching is necessarily good language teaching. I hope to show you, by way of examples from French immersion teaching, some ways in which typical content teaching is inadequate as a second language learning environment. And again, by means of examples, I hope to suggest ways in which content teaching can be manipulated and complemented to enhance its language learning potential.

There are many classrooms in Canada and elsewhere where the learning of content and the learning of a second language are both program goals. There is at least one major assumption about content teaching that is current in second language theory and pedagogical practices today. The assumption is -- because content teaching is considered communicative language teaching par excellence -- that through content teaching, second language learning will be enhanced.

This was certainly one of the assumptions underlying the initiation of French immersion programs. And just as it has come to be recognized in English content classes that learners of non-English backgrounds need the support of ESL classes, so in French immersion classes, the French language arts component is seen to support the language learning of the content class. What goes on in the content class, and the relationship of the language arts component to it, is the focus of this paper.

My guess is that most of us here accept the assumption that second language learning will be enhanced through content learning. However, there are pockets of evidence to suggest that such an assumption may be unwarranted.

Let us look, for a minute, at the results of an experiment carried out in Hong Kong (Ho, 1985). The first language of the students in the experiment was Cantonese and they had had English as a subject for six years and had been using textbooks written in English in all subjects in the previous year. The students, at the time of the experiment, were in grade 8, and consisted of the top-performing students in their school. The heart of the experiment was that for five months some of these grade 8 students were taught 60% of their curriculum totally in English, while other grade 8 students in the same school were taught all their curriculum in Cantonese. Students were randomly assigned to the English-instructed class or to the Cantonese-instructed class. The same teachers taught in both classes so that, for example, a teacher who taught Science in English to the English-instructed class taught the same lesson the same day to the other students in Cantonese.

Unfortunately, the write-up of the study tells us very little about the substance of the tests or the criteria used for evaluating English language performance, but it does tell us that at the end of the fifth month, there were no differences between the two groups in their performance on the English language tests which were given. In other words, five months of instruction in English using the content from a variety of academic subjects did not enhance the learning of English for these grade 8 Hong Kong students.

There are many possible explanations for this finding. For example, none of the teachers were native speakers of English. However, this is not atypical of ESL teachers in many parts of the world; and one would still expect some modest difference in English language performance between the two groups to be found. The explanation I consider most likely concerns the methodology of the presentation of the content.

Of course, I don't know precisely what methodology the teachers in the Hong Kong experiment were actually using, but I'm willing to make guesses. My guesses are based on what has been observed in typical content classrooms in the United States (Goodlad, 198X), and what I have observed in typical French immersion classrooms in Canada. The methodology is straightforward: teachers work through a content lesson by asking a lot of questions about something they have presented before, or that the students have read before. The teachers ask questions with particular answers in mind; students' responses are usually fairly short and to the point. This back and forth between teachers and their students is rapid and lively as students' hands bob up and down. Teachers tend to correct errors of content, and occasionally correct errors of syntax, morphology or pronunciation. Diversions from the main theme of the lesson arising from personal experiences or insights tend not to occur. Written seat-work may be assigned where the students answer more questions, or fill in blanks. This methodology, I would argue, leaves a lot to be desired from a language learning point of view.

Let's consider an actual lesson in more depth -- a few segments taken from a history lesson. The segments are taken from a grade 6 French immersion class, which, I've translated into English.

You will notice -- when you see the transcripts -- that the brief description I have just given you of typical content classrooms is evident. The teacher explains or summarizes facts and asks questions. The students reply with a word or short sentence. The teacher keeps them 'on-target', content-wise.

The three examples you will see represent only a tiny portion of data we have collected in a recent study. The study involved observing and tape-recording the full school day of nine grade 3 immersion classes and ten grade 6 immersion classes in Ontario schools. We have transcribed the tape-recordings and have begun to analyze the transcripts from a variety of perspectives (Swain and Carroll, 1987). One of the things we have looked at is the frequency and length of student talk in these teacher-fronted lessons.

Each student turn in each of the classes was categorized according to its length. They were categorized as 'minimal', 'phrase', 'clause' or 'sustained' in length. Minimal length refers to turns of one or two words in length. Phrase length refers to turns consisting of an adverbial phrase, a nominal phrase or a verb phrase; and clause length refers to a turn consisting of one clause. Any student turn which was longer than a clause was categorized as sustained talk.

The results indicate that there are, on the average, about two student turns per minute. In grade 6, about 44% of those are of minimal length. Only 18% of student turns are sustained in length. Those include occasions when students read aloud. When those occasions are subtracted, then it turns out that only about 14% of the times that students talk in teacher-fronted activities are their utterances longer than a clause. The figures are not much different for the grade 3 classes. As I will argue shortly, opportunities to produce sustained output in the second language are crucial to the second language learning process. Sustained talk provides both opportunities for variety and complexity of language use, and it forces the learner to pay attention to how content is expressed. This suggests that at least some portion of content lessons need to be structured in different ways in order to permit more opportunities for the sustained use of language by students.

Now, let us consider the excerpts from the history lesson taken from a grade 6 immersion class. The lesson is about the Antilles in 1796 -- what it was like then and the sorts of things that were influencing life at that time. Before we actually look at the excerpts, let me ask you to consider two questions. First, what do you think will be the most common tense used by the teacher -- past, present or future? Secondly, as a language teacher, what would be one reason to teach an historical theme?

I assume that you answered 'the past tense' to both of these questions. Now, let's see what happens when language is used authentically in the content classroom.

(1) T: It (Europe) didn't have sugar cane. Why didn't they have sugar cane? Mary?
S: It's too cold.
T: It's too cold. Another word for 'the weather'?
S: The climate is not good.

(2) T: What do you think? How did these plantations influence life in the Antilles? How do you think that these plantations...are going...uhm to change...life in the Antilles?

(3) T: These people are going to sell their sugar...rum... molasses...brown sugar. They are going to make money. With the money, they are going to buy clothes, furniture...horses...carriages...all that they want and they are going to bring back to the Antilles ...one imports...the Antilles imports...Now I want to

go back to what John was saying because I thought that
that was what he was trying to explain to me.
How is it going to change life in the Antilles?
S: Modernize.
.
.
.
T: OK. We are going to import modern objects...to the
Antilles. OK, it's one way that that's going to
influence things. Another...Is there another way of influ--
How are the plantations going to influence life in the
Antilles?
S: All the slaves and all the different cultures who work
on the uhm XXX.
T: Yes! You have these huge plantations...you certainly
are going to have some cultures and customs that are...
S: Different.
T: Are going to mix together.

Example (1) illustrates one of the teacher's relatively infrequent uses of the past tense in this history lesson. Notice that the student answers in the present tense. The teacher indicates acceptance by her repetition of the phrase, and concentrates on content by asking for a word that will, in her estimation, improve the response.

In the second example, we see the teacher switch from past tense usage in 'How did these plantations influence life in the Antilles?' to future tense usage 'How do you think that these plantations...are going...uhm to change...life in the Antilles?' Use of the 'immediate future', that is, the use of the verb 'to go' plus verb to signal action that is just about to happen, appears to be one of this teacher's favoured strategies in this lesson. Example (3) is illustrative.

These examples illustrate the conflict that arises between teaching content and teaching language. What the teacher has done by her use of the 'immediate future' is superb from a content teaching point of view. Its use has brought the distant past into the lives of the children, got them involved, and undoubtedly helped them to understand the social and economic principles which this historical unit was intended to demonstrate. However, as a language lesson, these examples illustrate several problems -- problems which may arise in any instructional setting based on authentic communication; problems which arise at the interface of language and content teaching.

First, the focus is entirely meaning-oriented. This is, of course, precisely what Krashen (198X) has argued is needed for second language acquisition to occur. He has argued that what learners should do is 'go for meaning'. But, if students are to actually acquire a second language by 'going for meaning', then they have to be engaging in some way in some sort of form-function analysis. That is, they will have to be paying attention to the form of the utterance as it is used to express the meaning they are extracting.

However, as Krashen (1982), himself suggests, "In many cases, we do not utilize syntax in understanding -- we often get the message with a combination of vocabulary, or lexical information plus extra-linguistic information". (p. 66). In other words, it is possible to comprehend input -- to get the message -- without a syntactic or, I would add, a morphological analysis of that input. What appears to occur is 'selective listening' (VanPatten, 1984a,b).

Awareness of such selective attention is illustrated in the part of an interview with an ESL speaker shown on the overhead (Wenden, 1983).

Q: Are you comfortable with him (the boss)?
A: Yes, he speaks slowly, more slowly than others, so it's easier for me.
Q: Do you ever notice how he says things?
A: When doing business, I don't consider grammar. Mostly I try to get the meaning. It's not necessary to catch all the words. (Wenden, 1983: 6)

Other kinds of evidence for selective listening exist. In one study of adult learners of Spanish, VanPatten (1983, 1984) isolated instances where learners apparently ignore how something was said to them. The example on the overhead is one such instance.

Q: Cómo están ellos? (How are they?)
A: Son contento. (They're happy.)
Q: Y ellos, cómo estan? (And how are they?)
A: Son contento tambien. (They're happy too.)
(VanPatten, 1984a: 92)

The learner appears not to have attended to the use of the correct copula 'estar' in the interviewer's first question and produced an utterance in which the wrong copula 'ser' was used. In the next question, therefore, the interviewer moved the copula to a more salient position -- the sentence final position -- but the learner still did not attend to how the interviewer phrased the question.

We have many similar examples from interviews conducted with French immersion students. The next example on the overhead is illustrative:

Q: Et qu'est-ce que tu ferais si tu gagnais la
   loterie? Si tu gagnais d'argent?
A: Je vais mettre dans la banque... (Harley and Swain, 1977: 41)

Here the question is asked using the conditional, and the student responds by using the immediate future form.

As VanPatten indicates, there are occasional reports on selective listening throughout the second language acquisition literature. What they all have in common is that "selective listening seems to involve concentrated focus on informational content and not necessarily on how that content was delivered." (1984b: 91) Additionally, linguistic literature on discourse argues for the notion of a fuzzy, open, non-deterministic syntactic parsing strategy that is used for comprehending discourse but that would be inadequate for producing it (see Clark and Clark, 1977, van Dijk and Kintsch (1983), Genesee (1986)). Thus, it may be that any grammatical processing involved in comprehension may be quite different from the closed logical system of rules required to produce a grammatical utterance. In other words, we can understand discourse without precise syntactic and morphological knowledge, but we cannot produce it accurately without precise syntactic and morphological knowledge.

Given that this is the case, then one role of the teacher becomes fairly evident: to help learners undertake the sort of form-function analysis needed to be effective communicators in their second language. This does NOT imply teaching rules, although it may well be an effective strategy for some aspects of language and for some learners.

What it does imply is that input that will help learners focus their attention on particular form-functional relationships is essential. This point, by the way, is the basis of the original title of this talk. Providing relevant input will necessarily be contrived: one has first to identify the area of focus, and then contrive contexts in which its use is natural. This will most certainly involve conscious reflection about the relationship between language form and content. I will return to this point shortly.

An equally important way to help learners focus their attention on particular form-functional relationships is to require them to produce language. If, as has been suggested, learners do not need precise syntactic and morphological knowledge to understand the gist of language input, but do need such knowledge for accurate production, then it will be by requiring students to produce that they will become aware of their grammatical needs. Their language production will have to be at more than a phrase or clause level if they are to learn the mechanisms for coherent and accurate discourse.

The second problem illustrated by the history lesson excerpts is closely related to the first -- that of the focus being entirely meaning-oriented. In concentrating entirely on meaning, teachers frequently provide learners with inconsistent and possibly random information about their target language use. IF the students are engaging in any sort of form-function analysis while listening to their teacher, consider the message relayed to them based on the first two history lesson examples. The message -- the hidden grammar lesson for the students -- is that past tense, the immediate future and the present tense are interchangeable. In example (1), a student's response in the present tense to a question asked about the past is accepted -- in fact repeated by the teacher, and, in example (2), the teacher switches from past to future within the same context.

As I mentioned earlier, these examples represent only a tiny portion of the data we have collected in an observational study of grade 3 and grade 6 immersion classes. One analysis we have carried out of these data involves the classification and counting of surface level grammatical errors made by the grade 6 immersion students as they interacted with their teachers. For each error, we noted whether the teacher corrected it. We counted both implicit and explicit instances of correction. Our findings show that only 19% of the grammatical errors students made were corrected, while the remainder were ignored by the teachers. The pattern of correction appears to be determined as much by an 'irritation' factor as by any consistent pedagogical or linguistic factors.

The solution is most definitely not to correct every error each time one is made. There is no research evidence to suggest that such a procedure would be effective, and it would certainly impede the flow of communication. What the most effective correction strateges might be is not clear. Again, I'll return to this point shortly.

The third problem, and one which cannot be inferred from any one individual example, is that what the students hear -- the input they receive -- may be functionally restricted. Certain uses of language may simply not naturally occur -- or may occur fairly infrequently -- in the classroom setting. When the main source of second language input is the classroom, this problem is particularly serious.

Let me give you two examples of what I mean by 'functionally restricted' uses of language. Both examples come from our observational study of French immersion classes. I would ask you to think of possible examples from your own ESL classroom teaching. This is not an easy task. The difficulty in doing so, is that you need to think about what is NOT there, not about what IS there. Furthermore, intuitions about one's own language use are frequently inaccurate. One way to start thinking about your own language use in class is to tape yourself teaching for a day. Later, you can listen and re-listen to it from a variety of different perspectives.

By 'functionally restricted', I mean that the full functional range of the linguistic item of focus is not used, or is infrequently used. One example is the use of 'vous' and 'tu' by French immersion teachers. We decided to look at this because we found that in tests of sociolinguistic performance, immersion students tended to overuse 'tu' in situations calling for the use of 'vous' -- situations such as making a req[illegible] to an adult. In other words, in formal contexts, 'vous' was underused by immersion students relative to native speakers of the same age. We thought the explanation for this might be linked to the input the students received in class.

The transcripts of the ten grade 6 classes were examined, and all instances of the teachers' uses of 'tu' and 'vous' were counted and classified according to the functions they served. The French pronouns 'tu' and 'vous' carry both grammatical and sociolinguistic information. A number distinction may be signalled by the use of the singular 'tu' versus plural 'vous'. A sociolinguistic distinction may be manifested in the familiar 'tu' versus the formal 'vous' which is a marker of respect or politeness.

If we look only at form, then no explanation for the immersion students' results emerge. That is to say, the transcripts reveal that 'tu' and 'vous' are used about equally often by immersion teachers -- on the average, each roughly about once a minute. When we look at function, however, the picture changes dramatically. It turns out that there are very few occurrences of 'vous' where it is used by teachers as a marker of politeness or deference -- less than one instance per class.

The second example of functionally restricted language use in immersion classes involves the use of verb tenses. We decided to look at the verb usage of teachers because correct use of non-present tenses is an area of continuing difficulty among immersion students. Among our findings is that students tend to overuse the passé composé, doing so in contexts where the imperfect should be used. Furthermore, the imperfect is rarely used with action verbs. We have also found that even at grade ten, immersion students correctly produce the conditional only a little more than half the time in obligatory contexts.

Our analysis of the teacher talk is not complete, but what we have found is, I think, interesting in light of the student performance results just mentioned. We began by looking at the frequency with which different verb forms were used by grade 6 immersion teachers. On average, over three-quarters of the verbs used by the grade 6 immersion teachers are in the present or imperative. The proportion of verbs in teacher talk in the past tense is approximately 15%; the future tense, 6%; and the conditional tense 3%. Of the verbs used in the past tense, about two-thirds are in the passé composé and one-third in the imperfect. The use of the imperfect was almost completely limited to the verbs avoir, ètre, faire and vouloir. Its use with action verbs was virtually non-existent. These figures, it seems to us, go a long way towards explaining the second language performance of the students.

To summarize to this point, there are many classrooms in which both the learning of academic content and a second language are -- or should be -- major goals. In traditional teaching of content, however, the language the teacher uses may be functionally restricted in certain ways, correction of content takes precedence over correction of form in order to preserve the communicative flow, correction of form that does occur is inconsistent in its message, and students' opportunities to engage in extended discourse are limited.

Solutions

What solutions can be offered?

Any solutions will have to have at least the following four characteristics. First, they will have to ensure that students obtain language input in its full functional range. Secondly, students must be given the opportunity to produce language in its full functional range. Thirdly, there will have to be a way of providing consistent feedback to learners about their language errors. Fourthly, any solution will have to help learners attend to their language weaknesses.

Perhaps it is best to begin by suggesting what are NOT solutions.

First, it is not a solution to suggest that teachers change their language use in teaching content. The language that is used is authentic -- it represents functionally motivated language. But it IS a solution to ask teachers to be aware of their language use so that they can engineer contexts which demand specific and otherwise infrequent uses of language.

As we saw in the history lesson, the teacher's use of the immediate future was strongly motivated on pedagogical grounds. Teaching the lesson using the past tense would have had the effect of distancing the events, and removing them from the immediate reality of the students. The solution is not to force language into content, but to explore content sufficiently so that language in its full range emerges. That takes time, and will only occur over a range of activities, topics and subjects.

Secondly, it is not a solution to correct all the language errors learners make during the content class. The flow of communication would quickly grind to a halt. The fact of the matter is that we do not know what error correction strategies might be most effective. There is surprisingly little research data on this important issue, and it is an area ripe for systematic study.

In the typical content class, with student talk and writing being as restricted as it is, students do not have to work at getting their meaning across accurately, coherently and appropriately. However, in the activities which I will discuss shortly, students are producing language for real audiences and a specific purpose. They are motivated to create their intended meaning precisely which involves grammatical accuracy, coherent discourse, and appropriate register.

Error correction derives its consistency from the stage in an activity in which it occurs. Students come to understand that there is a stage of 'spontaneous production' that will need to proceed through stages of revision and editing before it is 'publically presentable'. Through these stages of revision and editing, self and peer monitoring are as important as teacher feedback. Consistency in error correction also derives from the questions which generate it. That is to say, the questions which motivate error correction are along the lines of "Do you mean this, or do you mean that? It's not clear from what you've said. Or it's not clear from the way you've written this."

The needs of content learners as language learners argue for limiting the sort of content teaching observed in the history lesson, and increasing the opportunities for learners to hear and use language over a much wider range of activities within the topics and sujjects to be covered. Moving in this direction would be to recognize both the need of using language for content learning and of using content for language learning.

Examples

Examples of content teaching which take into account the needs of their learners as language learners can be found in some classes, schools, and Boards of Education across Canada. Here in Vancouver, one of the most interesting and one of the few systematic attempts to integrate content and language teaching for ESL learners is underway. The project team, under the leadership of Bernard Mohan and Margaret Early, will work with a group of teachers in the Vancouver School Board. Together they will prepare an activity-based content curriculum that will develop the language that is needed for academic content; that is, the language of description, sequence and choice; the language of classification, principles and evaluation. If you would like details about this project, I urge you to speak to Bernie or Margaret, both of whom are here at this conference.

Many other examples of content teaching adapted to the needs of second language learners can be found. The approaches taken have been two-pronged. The methodology of the content class itself has been modified to incorporate activities that demand extended use of written and oral language by students across a wide range of functions. Consider, for example, the history lesson on the Antilles. The teacher was trying to introduce the concepts of imports and exports; she was trying to show that life changed because of the flow of goods, and to indicate ways in which life changed as a result. A number of activities the students might undertake come to mind. For example, source books could be read, skits could be written and acted out, recipes using local Antilles' products could be located or concocted, descriptions of imported products could be written, advertisements could be created, and so on. Groups of students could research different stages of the importing/exporting process: finding buyers, preparing the product, packaging the product, managing staff, shipping, dock handling, and delivering. Each group could prepare descriptions of what needs to be done, identify problems and how to deal with them, write the needed letters, list the individuals that need to be contacted, and so on. Eventually each group could compare their findings with other groups.

But suppose that the teacher wanted to focus specifically on the use of the past tense. She might then ask students, for example, to imagine a situation where the goods ordered by a wealthy plantation owner had been paid for, but it was long after the agreed upon arrival date and the goods had not yet arrived. The task of each student is to write a letter to the importer inquiring about the order. Language such as "I ordered X on..., The order consisted of..." will be required. The letters could be sent to a classmate who must respond as the importer. Language such as "It was sent on..., It came back badly damaged, I received your payment only last week", and so on will be essential to complete the task. Of course, other tenses may be used, but the teacher may choose to focus only on the accurate use of the past tense in this particular exercise. Other aspects of language use could be built in. The tone of the plantation owner's letter could be discussed. Is the owner angry, business-like or friendly? What are the language forms that signal his or her state of mind? How should the importer respond? Should he or she respond differently depending on the plantation owner's tone? How can these differences be signalled through language?

The second approach to adapting content teaching to fit the needs of language learners has been to complement it with a language arts program. Here, the language implications of the content classroom activities can be explored in more depth. The Antilles letters could be followed up by other activities involving letter-writing. Letters to friends, letters to request information, letters to complain, letters to order goods, letters to invite, letters to refuse invitations, and so on, could be written. Focus on differences in style, the linguistic means by which politeness is expressed, the language of requests are matters which the students could explore. There is conscious reflection on the relationship between language form and meaning. Not only might the students write letters themselves to real individuals, but they might bring in letters sent to their parents, including the usual collection of junk mail for comment and analyses.

Recently I spent some time in Fair Oaks Elementary School, a school in Redwood City in California. The school is located in a low income, high minority, industrial area where Spanish is the primary language of most students. Many students, prior to the introduction of their current program three years ago, tested considerably below the fiftieth percentile on national tests in reading and language. Scores are now considerably higher, and the absenteeism rate is the lowest in the District.

The school describes itself as a bilingual, whole language school. In a brochure that the principal hands out to visitors, it says "Fair Oaks is a place where visitors can observe children...using reading and writing to learn about the world, using real books and writing real stories, discovering how to spell by writing and reading, critiquing each other's writings, [and] revising their work based on peer conferences...". The brochure also points out that Fair Oaks is a place where visitors can observe "teachers...who read aloud to students daily from a variety of books with rich language and complex ideas, [and]...whose instructional practices reflect their knowledge that...language skill development is embedded in genuine reading and writing, [and that] language is acquired through using it rather than practicing its separate parts...". School staff proudly point to the fact that no basal readers are found in the classrooms. Rather children are reading literature from published books and are creating their own texts.

The school's description of itself is no exageration. In fact, it seems to me to be somewhat of an understatement of the richness of language use that occurs in this school. Let me give you some specific examples of language arts activities I observed.

In a grade one class, children were working in small groups or individually. A couple of students were lined up to talk to the teacher. There seemed to be no need for discipline as the children were thoroughly absorbed in what they were doing. This state of organized calm was not created overnight, but by the gradual development of routines. The major activity while I was in the classroom was journal writing. Journal writing begins with each child writing a diary-type entry into their journal. It ends with a 'published' book. Children were at various stages in the process from journal entry to final publication. Some children were discussing it with their teacher or a fellow student. Some were expanding or correcting what they had written. Others were dictating their story to the teacher who wrote it correctly into a stapled set of pages. Yet others were illustrating their book, deciding who to dedicate their book to, reading their book to others, placing their book in the classroom library, or reading their classmates' books from the library.

At later grade levels, the same process was occurring but students were taking greater and greater responsibilty for the production of the language which appeared in their published books. In a grade 5 class, the process was written on the blackboard: choose a topic, write, conference, revise, edit, publish. The rule is that nothing gets published with errors. So when the students consider their work to be ready for publication -- when their limits on content and form are reached -- the teacher provides them with feedback about their remaining errors of form. This, then, is form corrected in a context created by students where the students, themselves, have signalled that they now need feedback.

In another class, I was shown through a radio broadcasting studio. Every Friday, for half an hour, a group of grade 5 and 6 students go on the air. The program consists, among other things, of news, stories, jokes, commercials, guest speakers, sensational citizens' awards, school and community announcements, and language arts projects. It is not difficult to see how the preparation of such a show incorporates the four characteristics required of a solution to the problems of traditional content teaching that I have outlined, and how easily the theme of any show could be related to any academic content being taught. The language which is needed for any particular radio show includes a variety of genres, and over time can encompass an endless range of language. Preparation of each script may involve reading newspapers, magazines, community flyers, cookbooks, joke books, content text-books, dictionaries, encyclopedias, and so forth. Knowledgeable resource people have to be decided on, contacted and interviewed. Recordings of the interviews can be made and transcribed. Notes have to be made, and these have to be translated into written texts. Scripts have to be written and perfected. Rehearsals -- to get it right -- have to take place.

On the day I visited, I was shown through the radio studio by the student secretary. She showed me the most recent letter they had received. It was in response to a letter they had written to Queen Elizabeth inviting her to be interviewed on their show. The Queen's letter was a perfect example of a formal letter of polite refusal. Although I did not see the letter that went to the Queen, I am sure that its form as well as its content was thoroughly debated and carefully produced.

The students involved are now providing training to students from other schools who wish to begin their own radio station.

In another classroom, I watched cross-age tutoring. Each grade 5 student had been paired with a kindergarten child. The grade 5 students had each chosen a book they thought their child would enjoy, and during their time together, their responsibility was to read the story to their kindergarten child and ask their child questions about the story. The older child wrote down the question, the kindergarten child wrote out a response, and the older child wrote the younger child's meaning underneath his or her response. After the kindergarten children had been returned back to their classroom, the grade 5 students returned to their desks to reflect individually on their experience by writing field notes. A teacher-led group discussion followed in which several of the students read aloud from their field notes. This was a daily event for the students. Through this activity, the students are given the opportunity for extended language use in both written and spoken form. Through their field notes, the older students learn to reflect on their child's language use and progress. Language becomes a focus of attention and analysis.

To summarize, I have tried to show that typical content teaching is not necessarily good second language teaching. Appropriately, content teaching focuses on comprehending meaning. However, what second language learners need is to focus on form-meaning relationships. Doing so is facilitated through the production of language, whether in written or spoken form. Because the typical question/answer sequence found in content classes tends to elicit short responses of minimal complexity from students, at least part of the content lesson needs to be substituted with activities which demand longer, more complex, and coherent language from the learners.

Focussing on form-meaning relationships is also facilitated through conscious reflection on the relationship between form and meaning in authentic language samples, and in their own language as students struggle to convey precisely their intended meaning. Students need to be guided through this process by engaging them in activities which have been contrived by the teachers to focus the learners' attention, and to naturally elicit particular uses of language.

Content teaching of the question/answer type is limited in the range of form/function relations it naturally brings with it. For this reason, it needs to be complemented with activities that make use of functions otherwise infrequently present. Again, the activities are contrived to ensure the authentic use of language forms.

And finally, content teaching with its focus on meaning, appears to provide unsystematic, possibly random feedback to learners about their language errors. It is not clear what strategies of error correction should be adopted. Certainly research has very little to say on the topic. The strategies advocated here, however, are to provide learners with the motivation to use language accurately, coherently and appropriately by writing for -- or speaking to -- real audiences. Preparation to do so will usually involve a process of revising and editing, and a commitment to an error-free final product. Error-free implies that learners have conveyed their intended meaning to their own -- and their teacher's -- satisfaction.

Thus, to facilitate second language learning, the typical question/answer sequence found in such content teaching could be largely substituted with carefully contrived activities, which bring into the classroom authentic language in its full functional range.

Chapter 6

FUNCTIONAL GRAMMAR IN FRENCH IMMERSION: A CLASSROOM EXPERIMENT[1]

Birgit Harley

1. PURPOSE AND RATIONALE

The purpose of this experimental study is to evaluate the impact on second language (L2) proficiency of a functional approach to the teaching of grammar in a French immersion context. Classroom materials used in the experiment concentrate on an area of French grammar known to be problematical for immersion students, and indeed for anglophone learners in general: namely, distinctions in meaning between two major past tenses, the imparfait and the passé composé. The context for the study is grade 6 early French immersion classrooms in the Province of Ontario.

The study has both empirical and theoretical foundations. It arises out of earlier research in immersion classrooms and theoretical debate concerning the input and interactional needs of classroom L2 learners. As an intervention in existing classroom practice, it can be justified in terms of recent approaches to L2 curriculum design.

1:1 French Proficiency Findings

The initial impetus for the experiment came from studies of the French proficiency of early immersion students at the grade 5 and 6 levels. These studies had indicated that after several years in their immersion program,[2] the L2 grammatical competence of these students still differed in systematic ways from that of native French speakers (Appendix C of Chapter 5; Harley and Swain 1978, 1984; Year 2 Report). This was despite notable strengths in global comprehension and in the ability to produce discourse that was generally coherent (though not necessarily grammatically accurate) in French (Swain and Lapkin 1982; Year 2 Report).

One area of French grammar where problems appeared particularly persistent was the verb system. For example, in an assessment of 69 grade 6 immersion students' grammatical competence in an oral interview, a relatively low accuracy rate (57% correct) for the use of verb forms in the future, imparfait, passé composé, and conditional was shown in the Year 2 Report (p. 161). Harley and Swain (1978), in an in-depth study of the oral use of French verbs by five randomly selected grade 5 students, found these learners using a largely shared verb system that was simpler than that of French and similar to English in certain fundamental respects. The students generally made basic distinctions in time, but they made little use of other more marked distinctions that were less regularly made in the French verb system or that were in some way incongruent with English. The continuing influence of the students' L1 in their second language development, which did not necessarily give rise to outright errors, has also been noted in other interlanguage studies (Harley and Swain 1984; Year 2 Report, pp. 65-70).

The focus of the present study on the use of the imparfait and the passé composé is motivated by the observation that immersion students by grades 5 and 6 might produce the forms of these two past tenses but without having grasped their different aspectual functions. At the grade 5 level, for example, Harley and Swain (1978) found in oral interviews that students appeared to be operating with one past tense per lexical verb, generally using the passé composé for actions, even in habitual and progressive contexts where the imparfait was required, and reserving the imparfait for stative verbs of inherent duration (e.g. *être*, *avoir*, *vouloir*). In written compositions at the grade 6 level, students have been observed to mix the imparfait and the passé composé indiscriminately in narrating a sequence of events, where the passé composé is the appropriate tense to use: e.g. Tout d'un coup, trois bandits aux visage cachés *sont entrés* dans la banque et *ont pris* leurs fusils puis *_disaient_... The confusion in use of these tenses may persist at higher grade levels. Even in grade 10, there is evidence that the contrasting functions of the imparfait and the passé composé have not yet been mastered by some early immersion students. The following exchange, for example, took place in an oral interview (Chapter 10) with a grade 10 student:

Interviewer: Qu'est-ce qu'elle faisait, ta mère, quand tu es partie pour l'école ce matin?
Student: Pardon, qu'est-ce qu'elle a dit?
Interviewer: Qu'est-ce qu'elle faisait?
Student: Oh qu'est-ce qu'elle faisait? Eh bien, elle *_a_ *réveillé* ma soeur et mon petit frère.
Interviewer: C'est ça qu'elle faisait quand tu es partie?
Student: Oui.

In the progressive context of *faisait* in the above example, the use of the imparfait was called for.

In short, based on immersion French proficiency findings, the imparfait/passé composé contrast in meaning appeared to be an appropriate topic on which to focus an experiment in the teaching of grammar from a functional perspective. Grade 6 was selected as an appropriate level at which to undertake the experiment for several reasons: (a) the students appeared by then to be familiar with the *forms* of the imparfait and the passé composé, (b) they were at an age when they might be able to benefit from a more explicit focus on grammar, and (c) proficiency findings appeared to indicate that students had not reached a point at which their grammatical competence might be considered fossilized (Harley and Swain 1984).

1:2 Input and Interaction Needs of Classroom L2 Learners

Recent theoretical statements concerning the input and interaction needs of classroom L2 learners have consistently emphasized the importance of authentic L2 use, in which interesting and relevant messages are conveyed. There is considerable controversy, however, on the issue of whether in promoting the goal of communicative competence, grammar needs to be taught.

According to Krashen (1982), for example, the teaching of grammar is of minimal benefit, in that conscious 'learning' cannot be converted into the central process of unconscious L2 'acquisition'. Students can use their learned competence only to 'Monitor', or edit, their L2 production under particular non-communicative circumstances and only if they are cognitively mature enough and have a particular type of Monitoring personality (Krashen 1982). In the immersion context, therefore, Krashen would see no useful purpose in teachers focussing on the L2 code, either by way of correction of errors or by grammar teaching. Indeed, in his view, immersion programs with their focus on subject matter teaching in the L2, are already producing the best possible results that can be achieved in a classroom context. Students become 'intermediates' who can use the language for communication and whose acquisition will continue on the outside "as they interact with and receive comprehensible input from native speakers" (Krashen 1984: 62-63). In Krashen's estimation: "Canadian immersion is not simply another successful language teaching programme — it may be the most successful programme ever recorded in the professional language-teaching literature" (Krashen 1984: 61). In this interpretation, then, any limitations in the French proficiency of immersion students have to do with inevitable limitations of L2 classrooms in general and, we may infer, the majority English-speaking social environment in which the immersion programs generally operate.

Krashen's whole-hearted endorsement of immersion programs is consistent with his theoretical position on the input and interaction needs of L2 students. In his view, the key to the success of immersion programs is the fact that 'comprehensible input' which he sees as "the only true cause of second language acquisition" (Krashen 1984: 61), has been provided via subject matter teaching. Comprehensible input is defined by Krashen (1982) in terms of i + 1, or input that is slightly in advance of a student's current level of competence and that is provided automatically in the natural, rough-tuned talk of teachers as they focus not on the L2 code but on conveying interesting and relevant messages in the L2. In keeping with his emphasis on the necessity and sufficiency of comprehensible input for L2 acquisition, Krashen de-emphasizes the utility of L2 production which he sees as a result, rather than a direct cause, of language acquisition. Thus: "the ability to speak a second language 'emerges' or develops on its own" (Krashen 1984: 61).

In essence, Krashen is claiming with respect to immersion that the students' ability to communicate in the L2 will not improve as a result either of attempting to fine-tune the input, or of teaching explicit grammar, or by increasing emphasis on student L2 production. The present experiment provides an empirical way of testing these assumptions, which have been questioned on theoretical grounds.

McLaughlin, Rossman and McLeod (1983), for example, elaborate a psychological model of L2 development which does not have recourse to the learning/acquisition distinction that Krashen makes. In their view, learners first apply controlled processes requiring active attention to language and other learning tasks. These processes may be conscious or unconscious and precede the use of automatic processes that no longer require active attention. Hence, it can be concluded, there is no reason to exclude grammar teaching from the L2 classroom, just as long as it is appropriate to the maturity level of the students. In fact, observations conducted in grade 6 immersion classrooms reveal that grammar is already being taught. However, the main emphasis in these grammar activities appears to be more on manipulating and categorizing forms than on relating forms to their meaningful use in communicative contexts (Chapter 5, Appendix C).

Others (e.g. Harley and Swain 1984; Schachter 1983; White 1985) have questioned the assumption that teachers' message-focussed input automatically provides the 'i + 1' that students need in order to make continued progress in the L2. Harley and Swain, for example, hypothesize that some problematic L2 features may be neither salient nor frequent in immersion teachers' natural, rough-tuned talk, that certain constructions that are in teacher talk are readily misanalysed along English L1 lines, and that teachers may be providing little feedback, or 'negative input' (Schachter 1984), that would alert students to errors in their production. Some evidence for Harley and Swain's hypothesis is provided by the interlanguage findings and by classroom observation data (Harley and Swain 1984; Chapter 5).

Krashen's assumption that comprehension and production skills in the L2 are equally well served by the simple provision of meaningful L2 input that the students understand has also been challenged. Swain (1985) and Rivers (1986), for example, argue that given the natural redundancy of language, one can often bypass grammar in comprehending, by processing the input at a semantic level and relying on one's knowledge of the world and situational cues. It becomes more important to analyse the L2 syntax when speaking or writing at any length, however, in order to make oneself clearly understood, or produce what Swain refers to as 'comprehensible output'. She hypothesizes that, with the reduction in time devoted to French schooling in the upper elementary grades of an early immersion program and an increasing school emphasis on reading and listening to subject matter information, immersion students may actually be getting progressively less opportunity to speak French in class, and hence have insufficient need to process the L2 at a syntactic level. What is needed, therefore, is more emphasis on comprehensible output in the immersion classroom. Swain suggests that if more contexts for individuals to engage in conversational interaction were provided in immersion classrooms, this could promote not only increased opportunities to negotiate for meaning (which, according to Long (1983) and others, should lead to an increase in comprehensible input) but would also provide more occasions during which students could be pushed to produce comprehensible output. Moreover, as Harley and Swain (1984) have argued, the students' opportunities for output may have to be planned from a linguistic point of view, in that the speech acts that occur naturally in the classroom context may provide little need or opportunity for students to produce some L2 forms. Findings from immersion classroom observations in grades 3 and 6 suggest that the students are producing relatively little sustained speech and that there are relatively few opportunities for students to engage as equal partners in conversational interaction. At the same time, there are few contexts in which the students need to use particular forms such as past tense verbs (see Chapter 5).

Together, these various views of the input and interaction needs of classroom L2 learners lead to the following specific hypothesis being tested in the present experiment:

> that the grammatical competence of immersion students, with respect to use of the imparfait and the passé composé, can be enhanced: (a) by providing focussed L2 input that promotes perception and comprehension of functional contrasts between these two verb tenses; and (b) by providing more opportunities for students to express these functions in the realization of interesting, motivating tasks.

1:3 L2 Curriculum Principles

A fundamental requirement of any curricular experiment representing an intervention in classroom practice is that it should be justifiable from an overall curriculum perspective. Assuming also in the present case that the concern is for fine-tuning rather than for a major reorientation of classrooom practice, the nature of the experiment should be such that it can readily be integrated with regular classroom activities.

Allen (1983, 1984) provides a theoretical perspective on L2 curriculum design which is clearly applicable. He argues for a flexible, variable-focus approach to L2 curriculum, incorporating three essential components: a 'structural-analytic' component in which the focus is on the formal features of language and medium-oriented practice, a 'functional-analytic' component in which the focus is on discourse and in which practice is both medium and message-oriented, and an 'experiential' component in which the focus is on the natural unanalyzed use of language (for personal, social, or academic goals) and message-oriented practice. These three instructional approaches can be considered complementary rather than mutually exclusive. Depending on the participants in, and circumstances of, a particular program, one of the three approaches will be primarily in focus. This does not, however, mean that another approach should not be in focus some of the time.

In the context of an immersion program, for example, which is identified by Allen (1983) as being primarily an experiential approach to L2 curriculum, there is room for reinforcement at the structural-analytic or functional-analytic levels when attention needs to be paid to formal or functional features of the L2. Allen (1984) provides examples of how a functional-analytic approach can be flexibly applied in an essentially experiential ESL context through the use of relatively self-contained units, or 'modules'. And Ullmann (forthcoming) describes the long-standing use of multi-media modules that are experientially focussed and which serve to add interesting and worthwhile content and opportunities for genuine communication to mainly form-focussed core French programs.

The present experiment is based on classroom materials that, in Allen's terms, have a functional-analytic focus, and that are attuned as closely as possible to the primarily experiential focus of an immersion program. The materials have characteristics of the modular approach, being designed as a set of activities that can be flexibly integrated with regular, ongoing immersion classroom practice. In general approach, therefore, the experiment is educationally justified from the perspective of L2 curriculum theory.

It should be emphasized at this point that the experiment is educational rather than psycholinguistic in character, in the sense that it is not designed to isolate and study a single variable with respect to its impact on L2 proficiency. Instead, it involves assessing the effect of a whole range of classroom activities which together combine a cluster of theoretically motivated characteristics, and which are put together in an educationally viable way (for a description of the materials see section 2:3 below).

2. DESIGN OF THE EXPERIMENT

The experimental study took place during the grade 6 year of immersion programs offered in three Ontario school boards, two of them in the Metropolitan Toronto area, and one in the Ottawa region. In a pre-test/immediate post-test/delayed post-test design, immersion students receiving the experimental treatment were compared with comparison students who were not exposed to the treatment and who continued with their immersion education as before.

2:1 The Sample

The sample for the study consisted of 319 grade 6 early total immersion students in twelve schools. In each of the three boards, there were four schools involved in the study; grade 6 immersion students in two of the four schools were included in the experimental group and grade 6 immersion students in the other two schools formed part of the comparison group. Altogether there were thus six experimental schools and six comparison schools, with the number of students divided approximately equally in the experimental and comparison groups. For the two boards in the Metropolitan area, the schools involved in the study were the only ones in which a comparable immersion program was being offered at the grade 6 level. In the Ottawa area board, the four schools were selected from a larger population on the basis of matching criteria outlined below. As a general principle, students in split grade 5/6 classes were omitted from the study. However, in one board in the Metropolitan Toronto area, there were not enough schools with 'straight' grade 6 immersion classes, and the grade 6 students from two schools with split 5/6 immersion classes were included in the sample. In one of these schools, the students formed part of the experimental group, and in the other school they were assigned to the comparison group. Table 1, p. 367, summarizes the distribution of experimental and comparison students in schools within each of the three boards.

In each board, experimental and comparison schools were matched as closely as possible with the help of board personnel. Matching criteria involved consideration of: the general socio-economic background of the immersion students attending each school, the years of teaching experience of the immersion teachers in each school, their sex, and their linguistic background. From information later obtained in a questionnaire addressed to the teachers themselves (see Appendix A, pp. 381 to 384), their general comparability was verified. Mean years of teaching experience was 9.8 for the teachers in the experimental group and 8.8 for those in the comparison group; estimated mean years of experience in teaching French in early immersion tended to be higher in the experimental group: 5.5 vs. 3.8 in the comparison group. Four teachers in each group had had additional experience teaching French to anglophones in other kinds of French-as-a-second-language programs (Core French, extended French, and late immersion), and three teachers in the comparison group had also taught French to francophones. There were three male teachers in the experimental group, two male teachers in the control group, and all others were female. With respect to linguistic background, four experimental and four comparison teachers were educated in French at all levels (primary, secondary, and post-secondary); for some of these teachers, post-secondary education was partly in English. In each group, there was one teacher who was educated in French at the primary level only, and in each group there was one teacher who was educated in French at the post-secondary level only. Almost all teachers indicated that they were primarily French-speaking or that they were balanced bilinguals. One teacher in the comparison group indicated that they were almost a balanced bilingual and one in the experimental group described themselves as essentially anglophone, with excellent skills in French.

On the same questionnaire, the teachers provided additional information about their grade 6 immersion classes. All the experimental teachers reported that their classes were average in French proficiency; of the comparison teachers, three reported likewise that their classes were of average proficiency, two reported a mixture of proficiency levels, one indicated that the class was above-average, and one teacher failed to respond. Schools differed as to the amount of time devoted to French instruction in grade 6. On average, according to the teachers' reports, the experimental classes spent 161 minutes per day in French, while the comparison classes spent a somewhat higher average of 186 minutes per day in French. Most of the classes were housed in traditional classrooms; however, two experimental teachers reported teaching their grade 6 classes in 'open' classrooms, and one comparison teacher's class was housed in a portable.

2:2 Procedures

Information about the study available via school board personnel to all schools prior to the experiment specified that the purpose of the study was to evaluate some newly developed French language materials focussing on an area of French grammar which was known to be problematical for immersion students. No mention was made of the focus of the study on the imparfait and the passé composé. Pre-testing in French of experimental and comparison students took place at the beginning of October soon after the school year had begun. So that the testing could take place during the same week in each of the three boards, a different tester was assigned to each board. All three testers were members of the Development of Bilingual Proficiency Project staff at O.I.S.E.; two were near-native speakers of French, and the third was a native French speaker.

The test materials consisted of a composition task, a cloze test, and an oral interview, all designed to assess the students' ability to make appropriate use of the imparfait and the passé composé (for details, see 2.4 below). The composition and cloze tasks were administered in that order to whole class groups in each school, while the tape-recorded oral interview was administered individually to a subsample of six students in each school. These students were chosen within certain restrictions: francophones (defined as students who spoke French regularly at home and had at least one francophone parent) were eliminated from the oral sample; from those remaining, the sample was randomly selected so that the number of males and females in the class was proportionately represented. Each of the three tests came in two forms, A and B, and in each school, half the sample received form A, and the other half received form B.

Immediately following the week of pre-testing, the teachers in the experimental schools took part in day-long workshops with project staff, during which the precise purpose of the study was explained, and the teachers were introduced to the classroom materials to be used over the next eight weeks. The teachers in the two Metropolitan Toronto area boards attended one workshop together at O.I.S.E. A project staff member then took part in a second workshop in Ottawa with the two teachers from the experimental schools in the third board. The teachers were asked to use particular activities in a given order and in designated weeks, and to fill out a brief evaluation sheet (see Appendix B, p. 385) for each week's activities, which also requested information about time spent using the materials. The confidentiality of the materials was stressed, to ensure that comparison teachers would not know the content.

During the eight-week period a project staff member made arrangements to visit the experimental teachers' classes in the Metropolitan Toronto area at times when they were using the materials. Two of the classes were observed on two occasions, and the other two classes once each.

After the eight weeks in which the materials were used in the experimental schools, experimental and comparison students were post-tested during the first week of December. Those who had been interviewed at the time of pre-testing were re-interviewed. For this set of post-tests, the test forms were reversed so that students who had received form A on a pre-test received form B on the relevant post-test, and vice versa. At this time, the teachers also received a questionnaire designed to verify background information about their classes and themselves (see Appendix A, pp. 381-384). The teachers who had been using the materials were allowed to keep them and were advised that they could continue to use them if they wished over the next few months, but that they were to remain confidential.

A follow-up round of delayed post-testing was then conducted in the experimental and comparison groups at the beginning of March to assess the long-term effect of the experiment. On this occasion, students received the same form of the tests that they had received in the pre-testing at the beginning of the study. All the teachers received a second questionnaire asking about their grammar teaching over the course of the year and, in the case of teachers who had used the materials, enquiring about whether they had re-used them in any way (see Appendix C, pp. 386-388). At this time, a complimentary copy of the materials was also made available to each of the teachers in the comparison group.

2:3 Description of Classroom Materials

The classroom materials used in the study were selected and adapted for use over the eight-week period from a bank of activities entitled Parlons du passé (Harley, Ullmann and Mackay 1985), which had been developed and piloted during the previous year in the context of a materials development project. This bank of activities, focussing on the imparfait and the passé composé, had been designed for flexible use as resource material not only in immersion programs but with older students in regular FSL classes and extended French programs. Some revisions were incorporated in the set of activities selected for the experimental study to permit more productive use of the target verb tenses by the immersion students.

The experimental version of the materials consisted of a single loose-leaf binder for the teacher in which each week's activities were clearly marked off with dividers. This division into weeks was not entirely in keeping with the resource principle inherent in the original bank of activities, but it was considered a necessary pre-planning move in the context of the experiment to ensure, as far as possible, that the experimental students would all have experienced the full range of activities. The teachers were free to use the relevant material at any time during the designated week. All the material, including the introduction, guidelines for the use of activities, etc. was provided exclusively in French. Each week's activities were presented in the same format to the teacher. An introductory paragraph, entitled Buts éducatifs, explained the educational goals of that week's activities. This was followed by a Préparation section, forewarning the teacher of any equipment or xeroxing of student material that would be needed. The third section, Présentation, provided detailed instructions as to how to proceed with the week's activity, or activities. This was followed by a teacher evaluation form for the week, and finally by any master copies and transparencies that would be required.

The introduction. A general introduction to the materials explained their purpose to the teachers, gave some linguistic background on the imparfait/passé composé distinction, and briefly outlined the content of each week's activities. The purpose of the materials was described in terms of the need, for L2 teaching in an immersion context, to bridge the gap between form-focussed grammar exercises and natural communicative subject matter teaching. It was argued that there was room for grammar teaching that provided focussed opportunities for the students to use specific forms in communicative contexts. Specifically stated goals were: (a) to establish the different meanings of the imparfait and the passé composé; (b) to integrate this grammar teaching with the teaching of worthwhile subject matter content and with the personal experience of students; and (c) to demonstrate an approach to grammar teaching in an immersion context that could be applied to other linguistic content.

The main functions of the imparfait and the passé composé were then outlined in non-technical terms with examples taken from the materials. It was explained that while the imparfait and the passé composé are alike in referring to past time, they contrast in expressing a different perspective on the past. The imparfait, it was pointed out, is normally used to express situations or events which are seen as unfinished or ongoing in the past and which are not perceived as having any specific beginning or end. Its essential incompleteness makes it the preferred tense for background description in the past, habitual or repeated past events, incomplete past actions or actions in progress that were interrupted. In contrast, it was noted, the main function of the passé composé is to refer to situations and events that are seen simply as having occurred at or by a given time. For example, the passé composé is typically used to refer to the main events in a narrative, or an event that interrupts another ongoing situation in the past. Adverbs, it was noted, often reinforce the different aspects of the imparfait and the passé composé. It was emphasized that the contrast in use of the two verb forms is essentially one of perspective rather than an intrinsic difference in situations or events themselves, and that a speaker or writer can sometimes exercise a choice depending on how an event happens to be perceived.

To help teachers see some of the difficulties faced by the immersion students, a brief sketch of major aspectual functions with French examples and their translation equivalents in English was provided (see Appendix D, p. 389), showing that in English, the distribution of simple past, past progressive, and present perfect forms is not congruent with the French system. Some examples of immersion students' errors were also supplied.

Following the linguistic background, the introduction to the materials concluded with the brief itemization of the week-by-week activities. These are now described in more detail.

Week 1: Le Loup-garou et le châle. The first week's unit focussed on a familiar kind of language arts activity to which immersion teachers and students are accustomed: namely, the reading of a story. This story, of culturally interesting content, consisted of an authentic French-Canadian legend about a werewolf, a theme that was revisited in later weeks. Originally told in oral form, the legend was adapted linguistically to provide the students with input clearly exemplifying the uses of the imparfait and the passé composé. The students' silent reading of the legend was to be preceded by a class discussion about French-Canadian oral literature traditions, about the supernatural in general, and about werewolves in particular, designed to draw on students' own knowledge of legends and to introduce cultural concepts that would enhance the students' comprehension of the legend. A series of comprehension questions, from which the teacher could select, followed the reading and were phrased in such a way as to elicit the use of the two tenses from the students. This was to enable the teacher to observe the extent to which students were able to use the imparfait and the passé composé appropriately. There were no activities during this first week designed specifically to focus the students' attention on the two target tenses.

Week 2: Passez-moi vos jumelles. In the second week, however, the legend was re-used in abbreviated form as the basis for a form- and function-focussed grammar lesson. Organized in two teams, each student in one of the teams had to underline the (20) instances of the imparfait used, while each student in the other team was required to underline the (20) instances of the passé composé. Pages had to be exchanged for correction across the teams, with points lost for missing or incorrect answers. New sheets were then to be handed out in which the abbreviated legend was divided into two columns (see Appendix E, p. 390), with clauses containing the imparfait ranged in one column and clauses containing the passé composé in the other. The students in small groups had to focus on one or other of the columns and try, initially by themselves, to figure out what it was about the underlined verbs in their column that meant that they were expressed in the imparfait, or in the case of the other column, in the passé composé. Depending on how well they could manage this linguistic detective task, the teacher could then help out with some rules of thumb (suggestions were provided for the teachers to use).

Week 3: Vive la différence. For this week, activities focussing students' attention on the incomplete/complete aspectual contrast between the imparfait and the passé composé were integrated with an opportunity for artistic expression. The students were first to be shown pairs of 'photos' representing past events, with captions beneath, that differed only with respect to the use of the two verb tenses: e.g. Il dégringolait l'escalier/ Il a dégringolé l'escalier. The actions displayed in the photos were either completed (passé composé) or were still in progress (imparfait). Students could then select three pairs of sentences from a list and make their own illustrations to show the distinction (e.g. Le second cheval a depassé le premier/ Le second cheval dépassait le premier; Marc soufflait les bougies de son gâteau/ Marc a soufflé les bougies de son gâteau).

Week 4: Proverbes. The activities for this week provided an opportunity to learn French proverbs along with opportunities for sustained oral production in referring to the past. In two activities, students had to apply proverbs from a list provided to past events. First, with the teacher's help, they were to relate 'Vouloir, c'est pouvoir' and 'On connaît ses amis au besoin' to the werewolf legend. Then, in a circle with the teacher, each student was to be invited to apply a proverb to a personal experience (e.g. Quand le chat n'est pas là, les souris dansent. Avoir les yeux plus grands que la panse).

Week 5: Qui et quoi? In a guessing game, a designated student (A), sent out into the corridor, had to guess what action was being mimed by another student (B) when student A knocked at the door. When student A entered, the class had to ask "Qu'est-ce qu'il/elle faisait quand tu as frappé à la porte?", providing an obligatory context for the use of the imparfait. This miming game was designed simultaneously to build students' action verb vocabulary in French, in that student B had to select what to mime from a given (or class-created) list of actions in particular domains (e.g. expression of feelings, exercise, household chores). A second activity involved an individual pencil-and-paper task. Using a picture representing a class preparing for a party, the students had to report on what was going on by filling the blanks in a text in which the imparfait was the appropriate tense. They then had to complete the story in a short written composition task, responding to questions about what happened next. The second part lent itself to use of the passé composé.

Week 6: Place à la fantaisie. In a return to the theme of the legend, the students in small groups working together, were invited to write new (comical) legends that included, for example, the werewolf together with a new character (e.g. Tintin, Alice, E.T.). A member of each group was then to read the legend to the class, and a jury would decide on the best legend, based on (a) interest and originality, and (b) frequency of appropriate use of the imparfait and the passé composé. The legends could be acted out if desired.

Week 7: Casse-tête. A short version of the werewolf legend was made into a jigsaw puzzle. Sentence pieces were provided in five envelopes (each containing an entire puzzle) and the students had to work in groups to put the pieces of their puzzle in the correct order, and fill in the blanks which had been left for the verbs. The infinitive of each verb was provided, and the required forms in context were, of course, the imparfait or the passé composé.

Weeks 7-8: Souvenirs de mon enfance. This series of activities was designed to draw on the students' personal experience in the creation of albums of childhood memories and in tape-recorded interviews. For their albums, students were invited to select two themes from a list of ten that were suggested, some of which provided opening phrases. For example:

> Un enfant unique.
>
> Quand tu demandes à tes parents ce qui te rendait 'unique' quand tu étais petit/e, qu'est-ce qu'ils répondent?
>
> (Commence ton récit avec les mots suivants: "Mes parents disent de moi que lorsque j'étais petit/e ...)

In this instance, a context was provided in which it is appropriate to use the imparfait. The album was to conclude with five questions that the student wished to be asked about his/her past. Another student would then interview the first student using these questions. The interview was to be recorded and the corrected albums and linguistically adequate tapes to be placed in a celebrity corner/listening centre for enjoyment by other members of the class.

Pour le suivi. In addition to the activities designed for use during the specified eight-week period, an environmental science unit, entitled À la recherche, involving research on wolves and other wild animals, was included in the material as a further example to teachers of how opportunities could be created in any subject area to integrate the teaching of grammatical functions with other worthwhile content. While the main focus of this substantial unit was on the wolf in reality, seen in relation to the wolf in fiction that had been the theme of some of the previous week's activities, there were several suggested activities in which contexts for use of the imparfait and/or the passé composé were created: e.g. in oral accounts of students' real-life experience of wolves, in the teacher's phrasing of questions about a movie on the life of wolves, in students' writing of stories in which they imagined themselves as an animal of their choice, and in research into extinct or prehistoric animals (references for French books were supplied). Teachers could select from these activities if they wished during the New Year and before the delayed post-testing.

The materials concluded with some suggestions to teachers as to how they might re-use, or build on, activities of the previous eight weeks.

2:4 French Proficiency Measures

In this section, the three kinds of measures used to assess the experimental and comparison students' ability to use the passé composé and the imparfait before and after the experiment are described.

Compositions. Composition topics (Form A = Aux voleurs!, Form B = Au secours!) which had previously been used with grade 6 immersion students in an earlier study (Year 2 Report, Chapter 2) were reused in the present experiment (see Appendix F, pp. 391-398). The students were given 15 minutes in which to write on the assigned narrative topic (Form A or Form B), which was provided with a brief opening that served to set a past time context. The instructions to the students, which were written in French, differed from the previous study in one respect: students were specifically instructed to use past tense in telling their stories. New scoring procedures were developed for the experimental study, with a focus on the ability to distinguish the forms and uses of the imparfait and the passé composé. A five-point scale was used by two native French-speaking raters to assess the accuracy in use of the two tenses (1 = high proportion of errors, 5 = no errors). For purposes of scoring, homophonous spelling errors in the passé composé, such as -ez or -er instead of -é, and agreement errors in both tenses were ignored, and the appropriate use of the literary passé simple instead of the passé composé was, of course, not penalized. If a student used only one of the target tenses — i.e. only the passé composé, or only the imparfait — appropriately in a composition, it was considered unscorable, since there was no indication that the student could actually distinguish between the two. (See Appendix F for details of scoring). All compositions were scored by both scorers and inter-rater reliability was high, with α s of .92 and .92 for the scoring of the two forms of the pre-test, .93 and .91 for the immediate post-test, and .91 and .94 for the delayed post-test (see Appendix F for further details).

Cloze tests. Cloze tests consisted of two narrative texts in French (Form A = Le vol and Form B = Le monstre) in which blanks replaced the use of the imparfait and the passé composé (see Appendix G, pp. 399-408). The narratives, based on longer, more complex French texts, had been adapted for use in the experiment by a francophone staff member of the project. Beneath each blank was written the needed verb in both the passé composé and the imparfait (in that order). The students, who had 25 minutes to complete the task, had to write the correct choice in the blank. There were 37 items in Form A and 38 in Form B. Five adults (two native speakers and three near-native speakers of French) were asked to complete the cloze tests at the beginning of the experiment in order to determine the validity of the items. There was full agreement concerning the correct choices except on one item in Form B, which was subsequently dropped from the analysis (see Appendix G). Items were scored as either correct or incorrect, with one point being given for each item that was correct. Failure to answer, choice of the incorrect alternative, or use of a mixed form (e.g. auxiliary with the imparfait, past participle with no auxiliary) were all considered incorrect choices.

Oral interviews. The interviews, in two forms — A and B, were structured so that questions provided contexts for the use of the imparfait and the passé composé (see Appendix H, pp. 409-415). For the purposes of scoring it was decided, owing to time restrictions, to focus only on those 'difficult' questions (1, 4 and 5) that required the use of the imparfait in association with action verbs (see section 1:1). In each form of the interview, questions 1 and 5 created contexts for the expression of past actions in progress, while question 4 created a context for expressing habitual past actions. In each form of the interview, question 1 was identical (Qu'est-ce que la classe faisait au moment où tu es parti pour venir ici?), but the content of the answer could be expected to differ on different occasions of test administration. Other interview questions differed in content in each form. Question 5 of Form A, for example, related to past actions in progress depicted on a beach scene, while in Form B the same kind of question related to a farm scene. Scoring, which was carried out by two near-native speakers of French, was based on transcripts of the responses to Questions 1, 4, and 5. For each of these responses, a count was made of the number of uses of the imparfait and the number of obligatory contexts for its use. (Further details of the scoring procedures are provided in Appendix H). Inter-rater reliabilities (correlations) on the three questions for Form A and Form B varied from .88 to .97 (see Appendix H for further details).

3. THE FINDINGS

The analysis of test results was based on the following overall scores for each of the three tests:

(a) compositions -- each student's rating on the 5-point scale of accuracy;

(b) cloze tests -- number of correct items out of a maximum of 38;

(c) oral interviews -- a percentage error score arrived at by dividing the total number of errors in use of the imparfait produced in response to Questions 1, 4 and 5 by the total number of obligatory contexts for the imparfait that were supplied.

Most of the statistical analyses of the test results are based on the school as the unit of analysis. For comparison purposes the main results of the study are also reported with student as the unit of analysis.

3:1 Preliminary Analyses

Based on the pre-test results, an item analysis was conducted on each of the forms of the cloze test to determine test reliability. With an N of 270 students, moderately good reliabilities were obtained: $\alpha = .610$ for Form A, $\alpha = .685$ for Form B.

Pre-test results on Forms A and B of each of the measures (composition, cloze, and oral interview) were compared in an analysis of variance in order to determine the relative difficulty of forms. As indicated in Table 2, p. 368, there were no significant differences between forms for any of the tests. In further analyses, test means in each case are therefore reported for the two forms combined.

A comparison of experimental and comparison students' scores on the pre-tests indicated that the comparison students tended to perform slightly better than the experimental students on all three measures (see Tables 3 and 4, pp. 369 and 370). In order to adjust for initial differences between the groups in assessing the effect of the experiment, analyses of covariance were conducted on post-test scores, using pre-test scores as a covariate.

3:2 Comparison of Experimental and Comparison Groups' Post-test Scores

An analysis of covariance of the immediate post-test results using the school as the unit of analysis (see Table 5, p.371) revealed that the experimental students were significantly superior to the comparison students on two out of three measures: the cloze test and the oral interview. There were no significant differences on the composition test. An analysis of covariance with student as the unit of analysis (see Table 6, p. 372) revealed the same pattern of results on all three tests.

Figure 1 (p. 379) presents the unadjusted means of the experimental and control students on the pre- and immediate post-tests in a graphic representation of their respective growth patterns during the course of the eight-week experiment. A steeper slope — i.e. more growth — may be noted for the experimental students on all three tests. Also of note is that the comparison students made virtually no headway on the cloze test between the pre-test and the immediate post-test. (For the pre-test mean scores, see Table 3, p. 369, for immediate post-test means, see Table 7, p. 373).

In order to determine whether a positive effect of the experiment would be maintained over the long term, a comparison was made of the scores of the experimental and comparison groups on the delayed post-tests, administered approximately three months after the immediate post-tests. Analyses of covariance were again used to control for initial differences in pre-test scores. Only those students were included in the analyses for whom data were also available on the immediate post-tests, in addition to their pre-test scores. The results of the analyses of covariance, presented in Tables 8 and 9, pp. 374 and 375, show that there were, in the long run, no significant differences between the experimental and comparison groups on any of the three measures. Delayed post-test means are provided in Table 10, p. 376.

A plot of the unadjusted means of the final sample on the three occasions of testing (see Figure 2, p. 380) shows graphically that the growth pattern for the experimental students is not the same between time 2 and time 3 (the immediate and delayed post-tests) as it was between time 1 and time 2 (the pre-tests and immediate post-tests). For each of the three tests the growth of the experimental students appears to be less rapid between time 2 and time 3 than it was between time 1 and time 2 when the experimental materials were in regular use. However, it should be noted that on each test there is nonetheless some continued growth for the experimental students between time 2 and time 3. In comparison, it can be seen that on the cloze and oral tests, the comparison students have made more rapid progress between time 2 and time 3 than they did between time 1 and time 2; at the same time their progress between time 2 and time 3 on the cloze and oral tests has been greater than that of the experimental students during the same period, resulting in more similar scores for the two groups on the delayed post-tests of these two measures. A different pattern may be observed for the composition test, where the comparison group has made virtually no progress from time 2 to time 3, but the experimental group has continued to make some progress.

In light of the fact that the experimental teachers tended to have more teaching experience in the early immersion context than did the comparison teachers, and given that the time devoted to French instruction tended to be higher in the comparison classes than in the experimental classes, step-wise multiple regression analyses were conducted to investigate the possible effects of these variables on the immediate and delayed post-test scores. With class means on the pre-tests entered first, followed by early immersion teaching experience, then daily amount of time devoted to French in grade 6, and finally treatment (experimental vs. control), results were consistent with those obtained in the analyses of covariance. No significant additional variance over and above that of pretest scores was accounted for by teaching experience or by amount of time devoted to French, while treatment was a significant predictor of immediate post-test scores on the cloze and oral measures.

3:3 Experimental Teachers' Evaluations and Use of Classroom Materials over Initial Eight Weeks

Table 11, p. 377, presents a summary of the experimental teachers' evaluations of the Parlons du passé materials, and the amount of time they reported spending on classroom use of the materials each week. These findings are derived from the weekly evaluation sheets that were provided for the teachers (Appendix B, p. 385). They show on a series of 5-point scales that, on average, teachers considered that student enjoyment of the weekly activities ranged from 3.0 to 4.3 depending on the unit concerned (1 = not at all, 5 = enormously); that difficulty of specific weeks' activities ranged from 2.7 to 3.8 (1 = very easy, 5 = very difficult); and that the educational goals of each unit had been realized at an average level of 3.5 to 4.7 (1 = not at all, 5 = very well) depending on the unit concerned. Teachers also indicated on their evaluation sheets the amount of time they had spent on individual units. This ranged from one to two hours on average. Among the six teachers, the total time that each reported spending on the materials over the eight-week period ranged from 9.3 hours to 16.5 hours, with a mean of 11.9 hours.

In addition to evaluating the materials on 5-point scales, the teachers were invited to comment each week on their experience with the materials. All six teachers regularly provided comments, which expanded on the ratings that they gave the materials. These comments are summarized in relation to each week's activities below.

Week 1: Le Loup-garou et le châle. The comments of teachers on the werewolf legend and associated activities were generally related to the thematic content, which was well received. Four teachers mentioned students' interest in, and enjoyment of, the legend. With respect to linguistic content, one teacher also noted that the legend had been easily understood. Two teachers mentioned that there had been animated discussion — in one case, of werewolf films that students had seen, and in another case, of a question about the stereotypic feminine role of the heroine in the legend. There were no comments with respect to students' use of the past tense in class discussions. On the question of how the materials were used, one teacher reported having read the story aloud to the students on one day, and then having students read it aloud the next day (the guidelines had suggested silent reading). Another teacher reported that they had had insufficient time to get a good discussion going.

Week 2: Passez-moi vos jumelles. This explicit grammar unit had been rated, on average, the most difficult of the activities by the teachers ($\bar{X}$ rating was 3.8, close to a 'quite high' level of difficulty). Most comments related to the linguistic content and to the built-in competitive, team-game approach. Teachers were divided as to whether the activity represented an interesting linguistic challenge or was too difficult for their students. Three teachers indicated complete satisfaction with the unit in their comments, and two teachers mentioned that the small group discussion of the functions of the imparfait and passé composé had gone well. Another referred approvingly to the motivating character of the team-game but noted that the passé composé team had had the more difficult task in identifying forms. The same teacher found the inductive approach to discovering the functions of the imparfait and passé composé in small groups too difficult for the students, and resorted instead to a more deductive approach involving the teacher and whole class together. This teacher was then pleased with the result: "J'ai passé beaucoup de temps à cette section mais je trouve que cela portera fruit." A second teacher also commented that the students had "great difficulty" in discovering the functions of the target verb forms, but without indicating if, or how, this had been resolved. Finally, two teachers noted, on an organizational level, that the system of counting points for the two teams was overly complicated.

Week 3: Vive la différence. This activity had been rated the easiest, on average, of all the activities (2.7 on the 5-point scale) and also the most successful in terms of enjoyment (4.3) and realization of educational goals (4.7). Of the five teachers who commented, two referred to the high level of enjoyment of the students, and several noted that the use of a pictorial medium was very helpful in promoting understanding of the functional contrast: "L'activité qui fait vraiment prendre conscience de la différence entre l'imparfait et le passé composé." One teacher mentioned that the students' illustrations were put on display. Another teacher mentioned having prepared a play in which the students had to use the imparfait and the passé composé.

Week 4: Proverbes. This week's activities received mixed reviews from the teachers. Three teachers had indicated in their ratings that students had only liked this activity 'a little', and two commented specifically that it was hard for students to understand the meaning of proverbs. On the other hand, one teacher had voted the students' enjoyment of this activity as enormous (5 on the rating scale), and another teacher indicated that the students understood the proverbs well and liked guessing what proverbs other students' stories illustrated. One teacher referred to the difficulty of listening to 30 stories one after the other, and changed the activity to a written one, with a few students reading their stories aloud each day. Another teacher suggested fishing for proverbs to add variety to the ones selected. None of the teachers commented on the linguistic content of the students' stories.

Week 5: Qui et Quoi. This unit consisted of two separate activities. With respect to the miming game, four teachers indicated that it had been a success, but one complained that it quickly became monotonous. One teacher noted that all but two students who took part in the miming had used the imparfait. Another teacher reported changing the game so that the student who had been designated to mime an action would fish it out of a bag and then ask either "Qu'est-ce j'ai fait?" or "Qu'est-ce que je faisais?" and other students would have to use the corresponding verb form in response. Note that this change eliminated the acting out of the action in progress interrupted by the knock at the door, which the original activity was designed to fulfil. The written activity, in which students completed a text based on a picture, raised the issue of correction of errors. One teacher asked whether the text should be corrected with the students. Another commented that the students were having problems with the imparfait endings and with conjugating reflexive and intransitive verbs in the passé composé. Another mentioned errors but not whether they had been discussed with the students. In two cases, there was no mention of written activities in the comments, and in one of these cases, only half an hour was reported to have been spent using the materials, suggesting that the written activity did not actually take place.

Week 6: Place à la fantaisie. A number of comments for this week focussed on linguistic content. The teacher who had noted formal problems with the imparfait and passé composé the preceding week, spent time this week revising conjugation of intransitive and reflexive verbs in the past before doing the week's activity, and expressed satisfaction with the students' performance: "Ça a marché mieux que la semaine dernière. Les élèves comprennent bien maintenant la distinction entre les deux temps-cibles." Four teachers indicated that the students had much enjoyed the creation of their own legends in groups and that the activity had been a success. In one instance, the teacher took over from the students the jury role for assessing use of the imparfait and the passé composé. Another teacher commented that two out of eight groups did not use the past tense in their legends and observed that this was a good success rate. Finally there was one teacher who gave the activity the lowest possible rating on student enjoyment and considered that the goals had been realized only 'a little'. The associated comment was "Rien de nouveau."

Week 7: Casse-tête. This group activity was favourably reviewed by the teachers. One teacher noted that some students inserted verbs in the blanks in the present tense, but that most used the passé composé and the imparfait, sometimes incorrectly. This teacher comments: "Au moins on se demandait si c'était L'imp(arfait) on le passé composé que l'on devait utiliser. On est conscient que les deux temps jouent un rôle respectif." Another teacher also comments on errors, but one teacher observes that there was an almost 100% success rate on the task, and that the group activity enabled weak students to share in success.

Weeks 7-8: Souvenirs de mon enfance. One teacher did not manage to carry out any of this week's activities and another indicated that there had been no time to tape interviews. A third teacher changed the activity because the students had already produced a similar kind of album as part of their English curriculum. Instead of producing an album, therefore, the students each chose one topic from the list provided, prepared and then presented it orally. According to the teacher, all the students used the passé composé and the imparfait, and students' enjoyment was very high. Another teacher, who also found student enjoyment very high, commented on how interesting the students' albums were, and noted that their taped interviews were very well done. In a final linguistic comment, one teacher noted that sometimes students used the passé composé but mis-spelled the past participle inflection as -ait.

General comments. Three teachers wrote additional general comments on the project. One indicated a desire to use the material again the following year "car je l'ai trouvé intéressant et excitant pour moi et les enfants." A second teacher suggested that a period of four weeks might have been preferable "pour la concentration et le suivi des activités." This teacher noted that the werewolf theme had aroused a great deal of interest among the students, that "Place à la fantaisie" had been the most popular and "Qui et Quoi" the least popular unit, and that "Vive la différence" was the best unit academically. A concluding observation made by the same teacher was that: "La méthode est bien variée et intéressante." Comments from the third teacher were offered following the delayed post-testing. This teacher explained that there had been little time to devote to further use of the materials in the New Year owing to various other program emphases. Summing up the perceived long-term effect of the project, this teacher observes that some students "éprouvent encore des difficultés avec les terminaisons, mais le choix du temps des verbes est plus exact. Ils ont trouvé les activités variées très intéressantes."

Classroom visits. Informal visits to four of the experimental classes were carried out in Week 4 of the experiment, when they were working on Proverbes. Two of the classes were then revisited in Week 5, and the miming activity of Qui et Quoi? was observed. The visits, which were made by a native French-speaking member of the project staff, were designed to check on questions such as whether the teachers were encountering any problems in using the materials, whether they were using the materials as indicated in the guidelines, how the students were reacting to the materials, and what teachers were doing with respect to student errors in production.

It was noted that in three of the four classes, teachers had changed the Proverbes activities so that students could prepare their oral presentations in advance. In one class, they had had several days to prepare (in writing). In all classes students appeared (and were reported by teachers) to be actively participating and enjoying the activities. It was noted that the teachers varied in the extent to which they corrected students' errors in verb tense/aspect. In one class, for example, the teacher corrected all misuses in past tense during the Proverbes activity, while in other classes, the teachers did not correct students as they slipped into the present tense in recounting their past experiences.

3:4 Results of Questionnaire 2: Focus on Grammar Teaching

This questionnaire, administered to experimental and comparison teachers at the time of delayed post-testing (see Appendix C, pp. 386-388), was designed partly to find out whether comparison teachers had, in their grammar activities, also focussed on the imparfait and the passé composé. This objective was in addition to the expressed intention of finding out, for purposes of further research, what grammatical features the teachers were concerned about. The imparfait and the passé composé simply appeared as items on a list of other grammatical categories that, it was anticipated, might be the focus of grammar teaching. The responses indicate that four of the six comparison teachers had spent the maximum hours indicated (5+ hours) on the passé composé and one had spent 3-4 hours. The sixth teacher indicated that grammar was only taught in that class on an individual student basis, when assignments were being corrected. Three of the comparison teachers also reported spending 5+ hours on the imparfait, and two reported spending 3-4 hours. For all five of the comparison teachers who indicated that they taught grammar, the passé composé appeared among the most frequently taught items, as did the imparfait for four of them. No attempt was made to determine whether the teachers were focusing on form only or on both function and form, since it was felt that any response would be biased by the nature of the question (teachers would feel bound to indicate that they were teaching both). These findings suggest, in any event, that the passé composé and imparfait were among the grammatical features most frequently being drawn to the attention of students in the comparison classes over the course of their grade 6 year. Teachers in the experimental classes, as might be expected, gave very similar responses to those of the comparison teachers with respect to the grammatical teaching of the imparfait and the passé composé, generally checking the '5+ hours' category.

An additional section was added to the questionnaires addressed to the experimental teachers, in order to determine whether they had made further use of the materials since December, and, if so, how much time they had spent on specific activities. Five teachers had re-used activities from the original eight-week period, and two used the À la recherche environmental science unit. Three teachers also reported using other activities as a point of departure. Table 12, p. 378, shows which activities the teachers reported using. Most frequent use was apparently made of Week 3: Vive la différence, followed by the Week 7 - 8: Souvenirs de mon enfance. All activities were reportedly used by at least one teacher either in their original form or as a point of departure for further activities. Number of hours spent was not clearly reported by all teachers. It appears to have varied from 1 3/4 hours to as much as 20-29 hours. The latter high estimate provided by one teacher, however, includes an estimated 10-15 hours spent on À la recherche, much of which may be carried out without reference to the past.

3:5 Descriptive Analysis of Difference Scores

With one exception, all the experimental schools improved from pre-test to immediate post-test on all measures (the exception being one school whose mean composition score on the immediate post-test was marginally below that of their pre-test score, resulting in a difference score of -.07). In contrast, four comparison schools lost ground between pre-tests and immediate post-tests (two schools on two measures, and two schools on one measure each). By the time of the delayed post-testing, another experimental school had lost ground relative to their pre-test scores on two measures (composition and cloze), and the experimental school that had done worse on the composition task on the immediate post-test continued to do worse than on the pre-test at the time of the delayed post-test. The remaining experimental schools all showed improvement on all measures from pre-tests to delayed post-tests. In the case of the comparison schools, three had lost ground from pre-tests to delayed post-tests, one school on two measures (composition and cloze), and two schools on one measure each — composition and cloze respectively.

Given the small number of experimental classes, it was not considered appropriate to do statistical analyses of class factors (such as time on materials, teacher evaluations of materials) in relation to post-test scores. However, it is perhaps relevant to note that the class that had been reported as spending the least amount of time using the materials in the initial eight weeks was one of the classes that lost ground at post-testing, and that the other class that lost ground was also below average in reported time spent on the materials. The class in which the teacher had been observed to be correcting students' verb errors in the Proverbes activity had, by the time of the delayed post-testing, the largest difference score (1.22) of any class in the study on the composition task and the highest mean composition score in absolute terms (3.82).

4. DISCUSSION

This experimental study started from the premise that the L2 grammatical competence of immersion students in grade 6 would benefit from a teaching approach that focussed the language input on a specific problem area and that gave the students more opportunity to practise producing relevant functional contrasts in the context of interesting, motivating tasks. The findings of the study indicate that there were some immediate benefits to the students who were exposed to an experimental treatment designed to reflect this approach. However, in the long run, the experimental students did not do significantly better than comparison students on the set of tests designed to measure their competence in a specific area of French grammar. These findings are now discussed in relation to the ongoing theoretical debate concerning L2 learners' input and interactional needs and in relation to information obtained during the experiment through teacher comments on the materials, questionnaires, and classroom observation.

The fact that there were no significant long-run differences between the experimental and comparison groups might be construed as support for Krashen's (1982, 1984) hypothesis that comprehensible input is all that is needed by the immersion students. This hypothesis rests on the assumption, however, that this is all that the comparison students were getting, an assumption which cannot be maintained in the face of previous classroom observations (Chapter 5, Appendix C), or in light of the questionnaire responses of the comparison teachers themselves, indicating that almost all their classes spent a considerable amount of time focussing on the code in grammatical activities (see section 3:5). Even the teacher who claimed not to teach grammar indicated that there were code-focussed discussions with students on an individual basis.

Instead, one reasonable interpretation of the findings is that for young students of the age concerned (11+ years), the experimental teaching approach attuned to their age level and interests served to accelerate L2 development in an area where more traditional methods of grammatical study could eventually also show results as students matured (i.e. on the delayed post-test). If this is a correct interpretation of the immediate, but not long-term, advantage of the experimental students, it underscores that age appropriateness is an important consideration in the design of grammar-teaching materials. In the present study, the grammatical domain selected for treatment was a renownedly difficult one for anglophones learning French (e.g. Darbelnet 1975, Moore 1981), yet it was possible to present it in a way which enabled 11-year-old immersion students to improve sooner than their non-experimental peers, while taking part in activities which they appeared in general to find interesting and enjoyable. At least some of the activities in the Parlons du passé materials, including Vive la différence which was generally considered the most useful for promoting awareness of the incomplete/complete aspectual contrast, are also appropriate for even younger students. In fact the Harley, Ullmann and Mackay (1985) materials were originally piloted in both grades 5 and 6 of early immersion, and teachers found them suitable at the grade 5 level. A similar kind of experiment might therefore be tried at a younger age level in immersion (with of course, some modifications to the academic nature of some of the activities), when the teaching of formal grammar is less prevalent. This would be the case, for example, in grade 3 or 4 (Chapter 5, Appendix C; Ireland, Gunnell and Santerre 1980). Starting a functional approach to grammar teaching younger might also be appropriate in view of an apparent tendency, noted by teachers, for immersion students' L2 development to slow down around grade 3 or 4. Students appear to develop strategies that are useful and sufficient for classroom communication, and that may become progressively more difficult to overcome (e.g. overusing the present tense to refer to the past).

The above interpretation assumes, however, that the experimental treatment that students received in the present study fully realized the potential of the teaching approach. This may be questioned on the basis of several kinds of information. In the first place, some of the teachers' comments suggest that there were organizational inadequacies in certain activities which needed to be overcome (e.g. aspects of the Passez-moi vos jumelles team game; ensuring that everyone could have a turn telling Proverbes stories). Secondly, the materials were designed to provide a wide variety of activities, and yet it was clear that some kinds of activities appealed to particular teachers (and presumably their classes) more than others. This suggests that a greater range of options in different areas might have led to greater benefits, and in fact, would have been more in keeping with the 'resource' nature of the original Parlons du passé bank of activities. Third, it appears from teachers' comments on some of their weekly evaluation sheets, that the content of the material did not necessarily promote a simultaneous focus on the code — even for the teachers. Not one teacher, for example, commented on the use of the imparfait and passé composé in the Proverbes activity. And from the classroom observations that were conducted during Week 4 of the study when the Proverbes activity was underway (see section 3:4) it appears that only one of four teachers was focussing on the verb forms that students were using. In three of the classes, students producing the present tense remained uncorrected. It is worth noting that the experimental class with a teacher who consistently corrected the students in such content-oriented activities was, of the four classes observed, the one with the largest difference score on the composition task from pre-test to either post-test. It appears that immersion teachers, when focussing on subject matter content in oral activities may not be aware of, or attend to, errors that students are making (Chapter 5; see also Chaudron 1977). Indeed, it may require considerable practice for teachers to attend to these errors when appropriate — i.e. particularly errors made in forms that are specifically being practised in meaningful oral activities. The instructions for the materials, and the single preparatory workshop, provided insufficient guidance to the immersion teachers with respect to the apparent need for 'negative input' (Schachter 1984).

It is clear from the comments made by the teachers that at least some of the experimental students continued to have trouble with accurate use of the imparfait and passé composé inflections in writing activities. This appears to be reflected in the lack of enhancement observed for the experimental students vis-à-vis the comparison students on the composition task, either on the immediate or the delayed post-tests. More activities to help students in the area of written distinctions from both a formal and functional perspective may be indicated. At the same time, it is possible that the scoring procedures for composition tasks were not sufficiently delicate to capture potential differences between the groups. Further more detailed analyses of the composition data, with recognition accorded to students who made appropriate use of the imparfait in association with actions, and not just with stative types of verbs such as avoir and être, could reveal further more subtle benefits for the experimental students.

Finally, it may be noted that the amount of time devoted to the use of the materials — an average of 11.9 hours per class during the initial eight weeks, with a few additional hours over the next three months — may not have been sufficient to show a sustained effect on the delayed post-test results. As previously noted (see section 3:5), it was two classes with relatively low amounts of time devoted to the materials that failed to make any progress from pre-test to delayed post-test on one or two of the specific measures.

5. CONCLUSION

The results of this experimental study suggest that the use by sensitized teachers of motivating, grammatically focussed materials attuned to the age level of immersion students in elementary school can accelerate grammatical development, even in an area of French grammar that poses subtle problems for advanced anglophone learners. The experiment did not attempt to isolate specific aspects of the input or interaction that may have had a key role in the results that were found, on the assumption that factors such as increased frequency and saliency in the input, greater and more focussed opportunities for output, goal-directed interaction in small group contexts, and appeal to students' meta-linguistic awareness will all have combined to produce the initial enhancing effect. It remains to be seen, based on further discussion with teachers, adjustment of materials in line with their comments, and the possible use of age-adapted material at earlier grades, whether a more long-lasting effect of such an approach to grammar teaching can be achieved in an early immersion context.

Footnotes

[1] I would like to acknowledge with gratitude the key creative role of Rebecca Ullmann and Lynda Mackay in designing the classroom activities on which this study is based.

[2] Beginning with a half-day kindergarten, students in an early total immersion program receive all their education in French during the initial years of schooling. Around grade 3 or 4, a period of English language arts is introduced, and by grades 5 and 6, the proportion of the day in French is reduced to about 50 per cent.

References

Allen, J.P.B. Functional-analytic course design and the variable focus curriculum. Revised version of a paper presented at the Eighteenth Annual TESOL Convention, Houston, March 1984.

Allen, J.P.B. A three-level curriculum model for second-language education. **The Canadian Modern Language Review, 40** (1), 1983, 23-43.

Allen J.P.B., Cummins, J., Mougeon, R. and Swain, M. **Development of Bilingual Proficiency: Second Year Report.** Toronto: Ontario Institute for Studies in Education, 1983 (mimeo).

Chaudron, C. Teachers' priorities in correcting learners' errors in French immersion classes. **Working Papers on Bilingualism,** 12, 1977, 21-44.

Darbelnet, J. L'imparfait, temps du passé inachevé. **Le Français dans le Monde, 114,** 1975, 6-9.

Harley, B. and Swain, M. The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper, and A.P.R. Howatt (Eds.) **Interlanguage.** Edinburgh: Edinburgh University Press, 1984.

Harley, B. and Swain, M. An analysis of the verb system used by young learners of French. **Interlanguage Studies Bulletin,** 3(1), 1978, 35-79.

Ireland, D., Gunnell, K. and Santerre, L. A study of the teaching and learning of aural/oral French in immersion classes. Ottawa: Ottawa Valley Centre, Ontario Institute for Studies in Education, 1980.

Krashen, S.D. **Principles and Practice in Second Language Acquisition.** Oxford: Pergamon Press, 1982.

Krashen, S.D. Immersion: Why it works and what it has taught us. In H. H. Stern (Ed.) **Language and Society** (Special issue), 12, 1984, 61-64.

Long, M. Native speaker/non-native speaker conversation in the second language classroom. In M.A. Clarke and J. Handscombe (Eds.) **On TESOL '82: Pacific Perspectives on Language Learning and Teaching.** Washington, D.C.: TESOL, 1983.

McLaughlin, B., Rossman, T., and McLeod, B. Second language learning: An information-processing perspective. **Language Learning, 33,** 1983, 135-158.

Moore, V. L'anglophone et l'apprentissage du passé en français. **Le Français dans le Monde, 160,** 1981, 39-44.

Rivers, W.M. Comprehension and production in interactive language teaching. **Modern Language Journal, 70**(1), 1986, 1-7.

Schachter, J. A universal input condition. In W. E. Rutherford (Ed.) **Language Universals and Second Language Acquisition.** Philadelphia: John Benjamins, 1984.

Swain, M. Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden (Eds.) **Input in Second Language Acquisition.** Rowley, Mass.: Newbury House, 1985.

Swain, M. and Lapkin, S. **Evaluating Bilingual Education: A Canadian Case Study.** Clevedon, Avon: Multilingual Matters, 1982.

Ullmann, R. A modular approach to second language teaching and learning. In S. Savignon and M. Berns (Eds.) **Initiatives in Communicative Language Teaching. Vol. 2.** Addison Wesley, forthcoming.

White, L. Against comprehensible input: The input hypothesis and the development of L2 competence. Montreal: McGill University, 1985 (mimeo).

Table 1

Distribution of Grade 6 Immersion Student Sample (N = 319)[x] at Time of Pre-tests

Board	Experimental group		Control group	
	School	N	School	N
A	a	34	c	24
	b	32	d	31
B	a	21[y]	c	10[y]
	b	30	d	35[z]
C	a	24	c	28
	b	25	d	25

[x] Note that the figures on this table do not necessarily indicate exact class sizes, since occasional students in some classes did not take the pre-tests (e.g. because of absence or failure to produce a parent permission letter).

[y] Grade 6 students from a split grade 5/6 class.

[z] Composed of two classes of 17 and 18 students each.

Table 2

ANOVA and Means of Pre-test Results on Forms A and B for Each of Composition, Cloze, and Oral Measures

Analysis of Variance

Measure	Source of variance	df	SS	MS	F	p
Composition	Forms	1	5.05	5.05	3.11	n.s.
	Error	255	413.87	1.62		
Cloze	Forms	1	0.2	0.2	0.01	n.s.
	Error	301	5973.1	19.8		
Oral	Forms	1	0.1021	0.1021	1.88	n.s.
	Error	61	3.3127	0.0543		

Means

		Form A			Form B	
Measure	n	Mean	sd	n	Mean	sd
Composition	132	3.06	1.27	125	2.78	1.28
Cloze	154	24.33	4.39	149	24.38	4.52
Oral	32	0.552	0.219	31	0.633	0.247

Table 3

ANOVA and Means of Experimental and Control Groups' Pre-test Results for Composition, Cloze, and Oral Measures, Using the School as the Unit of Analysis

Analysis of Variance

Measure	Source of variance	df	SS	MS	F	p
Composition	Groups	1	.0176	.0176	.097	n.s.
	Error	10	1.8183	.1818		
Cloze	Groups	1	.6912	.6912	.6329	n.s.
	Error	10	10.9212	1.0921		
Oral	Groups	1	.0025	.0025	.1020	n.s.
	Error	10	.2418	.0242		

Means

Measure	Experimental			Control		
	n	Mean	sd	n	Mean	sd
Composition	6	2.85	0.42	6	2.92	0.44
Cloze	6	24.01	0.01	6	24.49	1.08
Oral	6	.5826	.1348	6	.5539	.1737

Table 4

ANOVA and Means of Experimental and Control Groups' Pre-test Results for Composition, Cloze, and Oral Measures, Using the Student as the Unit of Analysis

Analysis of Variance

Measure	Source of variance	df	SS	MS	F	p
Composition	Groups	1	1.58	1.58	0.96	n.s.
	Error	255	417.34	1.64		
Cloze	Groups	1	45.32	45.32	2.30	n.s.
	Error	301	5927.89	19.69		
Oral	Groups	1	.0035	.0035	0.06	n.s.
	Error	61	3.4113	.0559		

Means

Measure	Experimental			Control		
	n	Mean	sd	n	Mean	sd
Composition	131	2.85	1.31	126	3.01	1.25
Cloze	160	23.99	4.14	143	24.76	4.75
Oral	33	0.5989	0.215	30	0.5839	0.258

Table 5

ANCOVA of Experimental and Control Groups' Immediate Post-test Results for Composition, Cloze, and Oral Measures, Using Pre-test Results as a Covariate and the School as the Unit of Analysis

Measure	Source of variance	df	SS	MS	F	p
Composition	Treatment	1	0.063	0.063	0.25	n.s.
	Error	9	2.244	0.249		
	Order	1	0.000	0.000	0.00	n.s.
	Order x Treatment	1	0.017	0.017	0.37	n.s.
	Error	9	0.425	0.047		
Cloze	Treatment	1	18.894	18.894	11.91	.01
	Error	9	14.279	1.587		
	Order	1	2.058	2.058	1.05	n.s.
	Order x Treatment	1	0.239	0.239	0.12	n.s.
	Error	9	17.598	1.955		
Oral	Treatment	1	0.0523	0.0523	7.81	.05
	Error	8	0.0535	0.0067		
	Order	1	0.0154	0.0154	1.85	n.s.
	Order x Treatment	1	0.0356	0.0356	4.29	n.s.
	Error	8	0.0663	0.0083		

Table 6

ANCOVA of Experimental and Control Groups' Immediate Post-test Results for Composition, Cloze, and Oral Measures Using Pre-test Results as a Covariate and the Student as the Unit of Analysis

Measure	Source of variance	df	SS	MS	F	p
Composition	Treatment	1	0.835	0.835	0.73	n.s.
	Order	1	0.227	0.227	0.20	n.s.
	Treatment x Order	1	0.251	0.251	0.22	n.s.
	Error	252	287.596	1.141		
Cloze	Treatment	1	72.066	72.066	4.80	.05
	Order	1	49.427	49.427	3.29	n.s.
	Treatment x Order	1	7.820	7.820	0.52	n.s.
	Error	298	4473.251	15.011		
Oral	Treatment	1	0.143	0.143	5.28	.05
	Order	1	0.035	0.035	1.29	n.s.
	Treatment x Order	1	0.105	0.105	3.87	n.s.
	Error	58	1.573	0.027		

Table 7

Mean Scores[a] for Experimental and Control Groups' Immediate Post-test Results on Composition, Cloze, and Oral Measures, Using the School as the Unit of Analysis (A), and Using the Student as the Unit of Analysis (B)

A. School as the Unit of Analysis

Measure	Experimental			Control		
	N(ab)	N(ba)	Mean	N(ab)	N(ba)	Mean
Composition	6	6	3.19	6	6	3.13
Cloze	6	6	25.50	6	6	24.43
Oral	6	6	0.4291	5	5	0.5261

B. Student as the Unit of Analysis

Measure	Experimental			Control		
	N(ab)	N(ba)	Mean	N(ab)	N(ba)	Mean
Composition	63	68	3.20	69	57	3.14
Cloze	82	78	25.43	72	71	24.93
Oral	17	16	0.4449	15	15	0.5260

[a] Each mean is the average of the mean scores for Order (ab) and Order (ba). The number of cases contributing to the mean for each Order is shown.

Table 8

ANCOVA of Experimental and Control Groups' Delayed Post-test Results for Composition, Cloze, and Oral Measures, Using Pre-test Results as a Covariate and the School as the Unit of Analysis

Measure	Source of variance	df	SS	MS	F	p
Composition	Treatment	1	0.281	0.281	0.68	n.s.
	Error	9	3.747	0.416		
	Order	1	0.038	0.038	0.34	n.s.
	Order x Treatment	1	0.228	0.228	2.00	n.s.
	Error	9	1.027	0.114		
Cloze	Treatment	1	12.288	12.288	3.33	n.s.
	Error	9	33.162	3.685		
	Order	1	2.728	2.728	2.74	n.s.
	Order x Treatment	1	0.035	0.035	0.04	n.s.
	Error	9	8.961	0.996		
Oral	Treatment	1	0.0078	0.0078	0.26	n.s.
	Error	8	0.2368	0.0296		
	Order	1	0.0580	0.0580	2.67	n.s.
	Order x Treatment	1	0.0018	0.0018	0.08	n.s.
	Error	8	0.1737	0.0217		

Table 9

ANCOVA of Experimental and Control Groups' Delayed Post-test Results for Composition, Cloze, and Oral Measures, Using Pre-test Results as a Covariate and the Student as the Unit of Analysis

Measure	Source of variance	df	SS	MS	F	p
Composition	Treatment	1	1.363	1.363	1.16	n.s.
	Order	1	0.717	0.717	0.61	n.s.
	Treatment x Order	1	5.343	5.343	4.55	.05[a]
	Error	215	252.596	1.175		
Cloze	Treatment	1	10.716	10.716	0.74	n.s.
	Order	1	16.154	16.154	1.12	n.s.
	Treatment x Order	1	0.028	0.028	0.00	n.s.
	Error	265	3811.948	14.385		
Oral	Treatment	1	0.006	0.006	0.14	n.s.
	Order	1	0.030	0.030	0.68	n.s.
	Treatment x Order	1	0.003	0.003	0.08	n.s.
	Error	49	2.135	0.044		

[a] No interpretation of this significant interaction was apparent.

Table 10

Mean Scores[a] for Experimental and Control Groups' Delayed Post-test Results on Composition, Cloze, and Oral Measures, Using the School as the Unit of Analysis (A), and Using the Student as the Unit of Analysis (B)

A. School as the Unit of Analysis

Measure	Experimental N(aba)	 N(bab)	 Mean	Control N(aba)	 N(bab)	 Mean
Composition	6	6	3.34	6	6	3.06
Cloze	6	6	25.88	6	6	25.40
Oral	6	6	0.3706	5	5	0.3989

B. Student as the Unit of Analysis

Measure	Experimental N(aba)	 N(bab)	 Mean	Control N(aba)	 N(bab)	 Mean
Composition	48	61	3.29	62	49	3.14
Cloze	71	70	25.63	66	63	25.77
Oral	15	15	0.3844	13	11	0.3909

[a] Each mean is the average of the mean scores for Order (aba) and Order (bab). The number of cases contributing to the mean for each Order is shown.

Table 11

Evaluation and Use of the Experimental Classroom Materials by Teachers

Week	Name of unit	$\bar{x}$ enjoyment (1 - 5)	$\bar{x}$ difficulty (1 - 5)	$\bar{x}$ realization of goals (1 - 5)	$\bar{x}$ hours of use
1	Le Loup-garou	3.7	3.3	3.7	1.6
2	Passez-moi vos jumelles	3.5	3.8	4.2	1.4
3	Vive la différence	4.3	2.7	4.7	1.2
4	Proverbes	3.0	3.3	3.8	1.4
5	Qui et quoi	3.3	3.3	4.0	1.5
6	Place à la fantaisie	3.3	3.2	3.5	1.7
7	Casse - tête	3.8	3.3	4.0	1.0
7+8	Souvenirs de mon enfance	4.0	3.4	4.0	2.0

Table 12

Further Use of Materials Reported by Experimental Teachers

	Number of teachers re-using material	Number of teachers using material as a point of departure[a]
Le Loup-garou et le châle	1	
Passez-moi vos jumelles	1	1
Vive la différence	3	2
Proverbes	2	
Qui et Quoi		1
Place à la fantaisie	1	1
Casse-tête	1	
Souvenirs de mon enfance	2	2
À la recherche		2

[a] Two teachers apparently interpreted this to mean the original use of the materials in the fall eight-week period. Their tallies are not included in this column, except for the À la recherche unit which one of the teachers reported using.

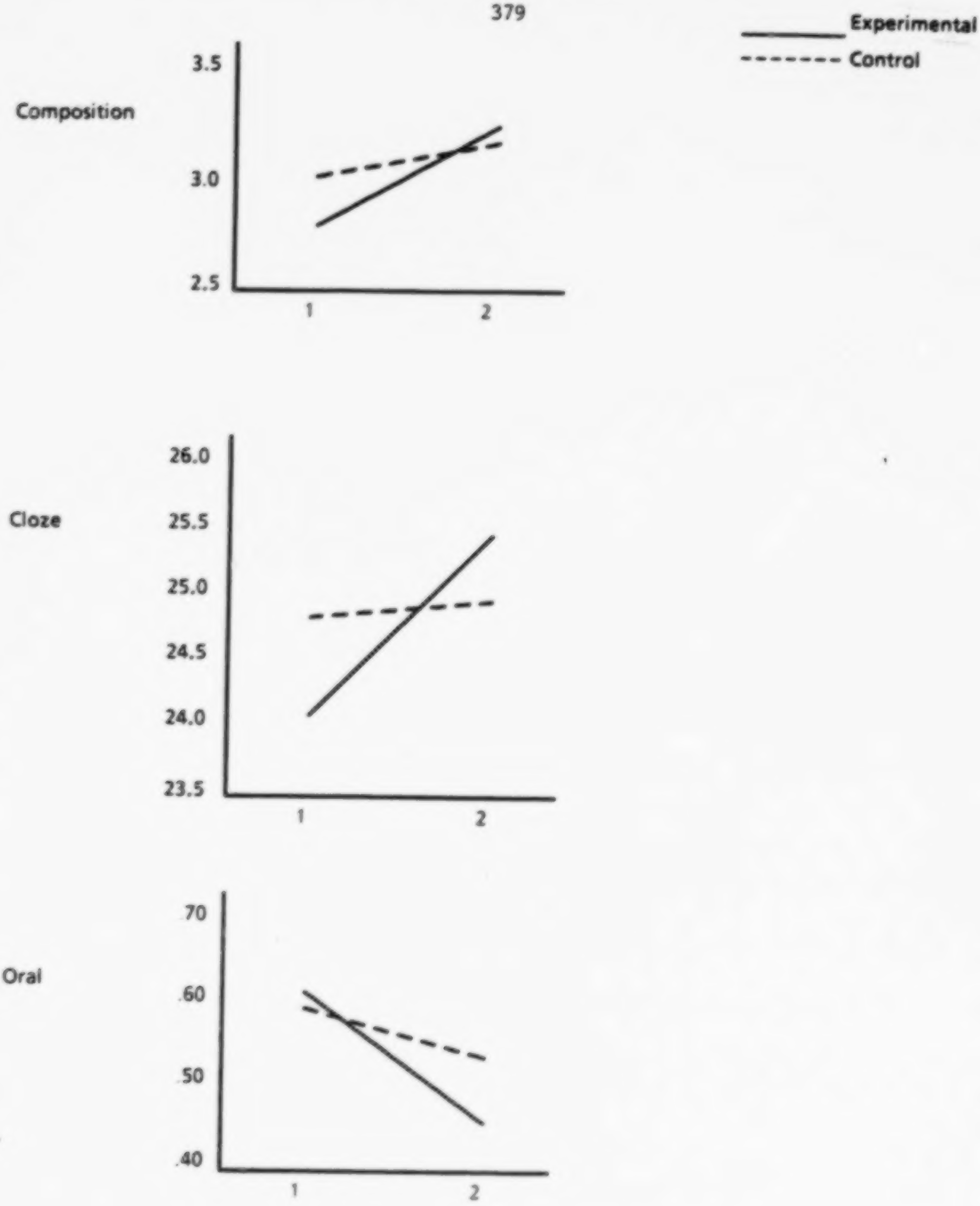


Figure 1

Unadjusted Means on Pre-test (1), Immediate Post-test (2) for Experimental and Control Groups on Each Measure, with the Student as the Unit of Analysis

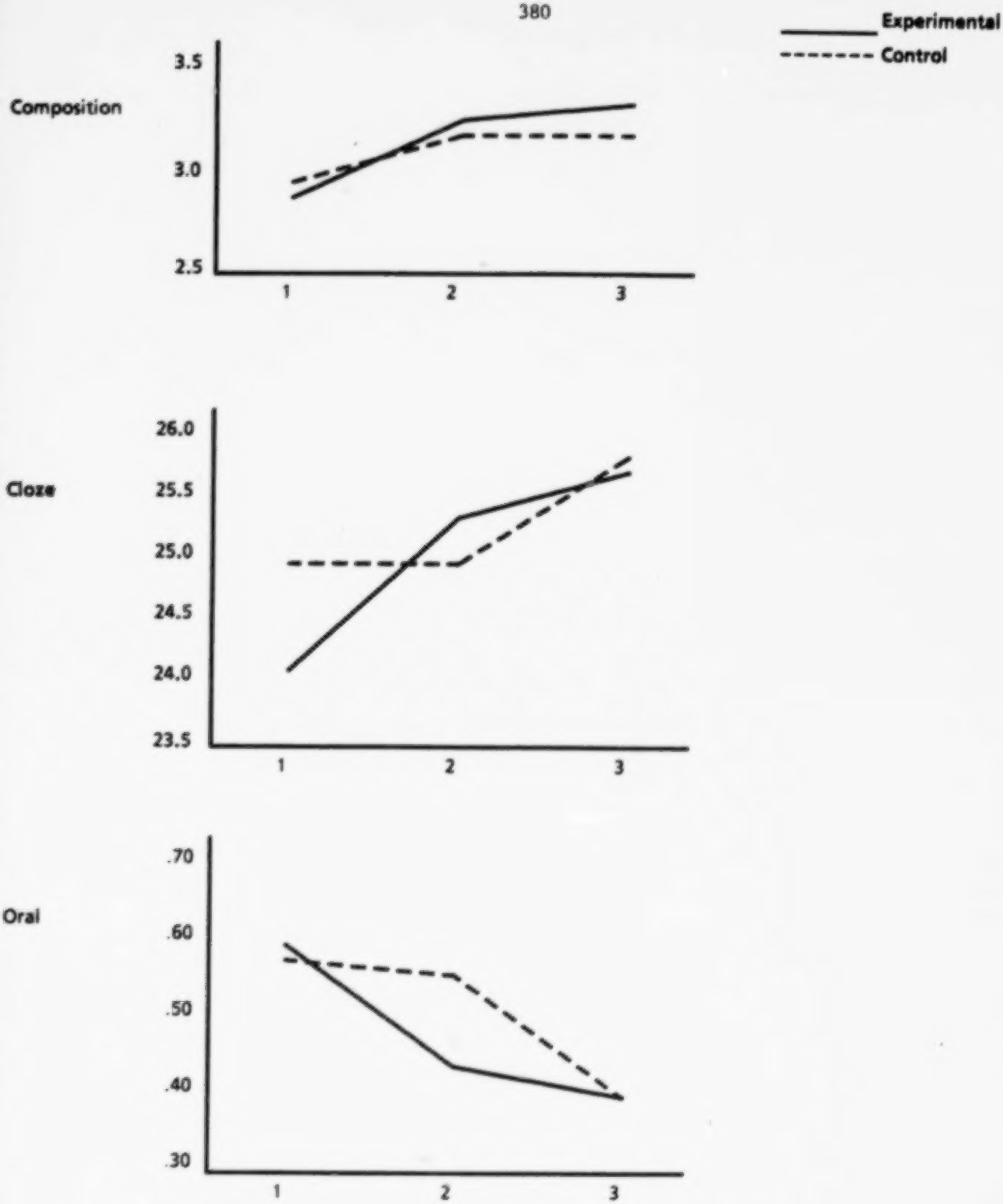


Figure 2

Unadjusted Means on Pre-test (1), Immediate Post-test (2), and Delayed Post-test (3) for Experimental and Control Groups on Each Measure, with the Student as the Unit of Analysis

Appendix A

Questionnaire à l'intention des enseignants du programme d' immersion en français de 6e année

A - Renseignements généraux

1. Quelle est votre expérience dans l'enseignement? Veuillez inscrire le nombre d'années dans la(les) case(s) appropriée(s).

Nombre total d'années d'enseignement ________________
Indiquez également cette année.

	K1-3	4-6	7-8	9-12	Nombre total d'années
Enseignement en anglais à des anglophones (programmes réguliers d' anglais)					
Enseignement en anglais à des anglophones (programme d'immersion totale en français					
Enseignement en français à des anglophones (programme de base, c.àd. une période de français par jour)					
Enseignement en français à des anglophones (programme intensif, c.àd. un cours de français plus une autre matière en français)					
Enseignement en français à des anglophones (programme d'immersion)					
Enseignement de l'anglais à des francophones					
Enseignement en français à des francophones					
Fonction particulière (bibliothécaire...) ________________					

2. Décrivez votre programme d'enseignement cette année. (Vous pouvez, par exemple, être titulaire d'une 6ème année - immersion - et, en plus, enseigner un programme de base à deux classes de 5ème année.)

____

____

____

____

B - Renseignements relatifs à votre programme

1. Dans quel type de classe enseignez-vous?

☐ décloisonnement (classe ouverte) ☐ classe traditionnelle ☐ bâtiment portatif

2. Votre école comporte-t-elle une aile ou une section française? ☐ oui ☐ non

3. Combien avez-vous d'élèves dans la classe qui fait partie de l'étude? ________________

4. Comment évalueriez-vous le niveau de vos élèves quant à leur connaissance du français?

☐ au-dessus de la moyenne

☐ moyen

☐ au-dessous de la moyenne

☐ hétérogène

5. Dans l'ensemble, votre classe est-elle un groupe homogène depuis l'entrée à l'école des élèves?

☐

ou

est-elle constituée d'élèves (de quelques élèves) différents chaque année?

☐

6. Approximativement combien de temps par jour passez-vous avec vos élèves de 6^{e} année?

________________________heures/minutes

7. Approximativement combien de temps pendant la journée vos élèves passent-ils à étudier:

en français ____________ h./min.

en anglais ____________ h./min.

8. Qui enseigne les différentes matières dans votre classe et en quelle langue?

	moi-même	enseignant d'une autre classe	enseignant particulier	anglais	français
étude de la langue française				-----	-----
étude de la langue anglaise				-----	-----
mathématiques					
sciences sociales					
sciences					
musique					
art					
éducation physique					
autre ____________					

C - Origines linguistiques

1. Quelle a été la langue habituelle d'enseignement au cours de votre scolarité?

primaire ____________________

secondaire ____________________

université ____________________

brevet d'enseignement (si vous l'avez obtenu hors de l'université)

____

2. Quelle langue utilise-t-on principalement chez vous?

☐ anglais

☐ français

☐ les deux

☐ ni l'une ni l'autre

3. Quelle langue utilisez-vous le plus lors de vos rapports sociaux?

☐ anglais

☐ français

☐ les deux

☐ ni l'une ni l'autre

4. Quelle langue utilisez-vous le plus lors de vos relations professionnelles quotidiennes?

☐ anglais

☐ français

☐ les deux

5. Aujourd'hui, comment vous décririez-vous linguistiquement?

☐ également compétent(e) (bilingue équilibré(e) en français et en anglais?

☐ essentiellement francophone avec connaissances approfondies de l'anglais,

☐ essentiellement anglophone avec connaissances approfondies du français.

☐ autre (spécifiez s.v.p.) ____________________

____

6. Quelle(s) autre(s) langue(s) parlez-vous bien? ____________________

Merci beaucoup de vous être donné la peine de remplir ce questionnaire pour nous

The Ontario Institute for Studies in Education

APPENDIX B

FEUILLE D'ÉVALUATION

Professeur: ______________________________

École: ______________________________

Numéro de la semaine: ____________________

Nom de l'unité: ___________________________

Faites un crochet (   ) dans la case appropriée:

1: En général, les étudiants ont aimé les activités de cette unité

pas du tout	un peu	assez	beaucoup	énormément
( )	( )	( )	( )	( )

2. En général, le niveau de difficulté de cette unité était

très élevé	assez élevé	moyen	bas	très bas
( )	( )	( )	( )	( )

3. Les buts éducatifs de cette unité ont été réalisés

pas du tout	un peu	assez bien	bien	très bien
( )	( )	( )	( )	( )

4. Combien de temps avez-vous mis à faire cette unité? ________ heures

Commentaires: ____

____

____

____

____

____

____

Appendix C

Questionnaire 2: Focus of Grammar Teaching

Questionnaire destiné aux enseignants de sixième année d'immersion précoce qui ont participé à l'étude expérimentale

Nous vous remercions beaucoup d'avoir bien voulu participer à cette étude expérimentale de matériel pédagogique. Le but de l'étude a été d'évaluer l'impact qu'a eu le matériel sur la compétence en français des élèves de 6^{e} année en immersion précoce. A cette fin nous sommes en train de comparer les classes qui ont reçu le matériel avec d'autres qui ne l'ont pas reçu. Nous vous ferons parvenir une copie du rapport quand il sera fini. Naturellement, l'identité des classes et des écoles restera entièrement confidentielle.

Nous voulons maintenant vous demander quelques questions à propos de vos priorités en ce qui concerne l'enseignement de la grammaire en sixième année. De plus nous recherchons votre avis sur la disponibilité de matériel pédagogique destiné à l'enseignement de cette grammaire. Ces renseignements nous aideront à savoir où concentrer nos recherches à l'avenir.

Au cours de l'année, sur quelles formes avez-vous concentré votre enseignement de la grammaire:

	Nombre total d'heures mis à l'enseignement en 6e année				Choix de matériel publié		
	Pas du tout	1 - 2 heures	3 - 4 heures	5 + heures	insuffisant	satisfaisant	excellent
a) Les pronoms personnels							
b) les adjectifs							
c) le genre							
d) les temps verbaux:							
le présent							
le passé composé							
l'imparfait							
le passé simple							
le futur simple							
le futur proche							
le conditionnel							
e) les verbes pronominaux							
f) les auxiliares							
g) les objets indirects							
h) autre(s) (spécifiez SVP) ________ ________ ________							

Merci beaucoup de votre coopération.

L'Institut d'Etudes Pédagogiques de l'Ontario
The Ontario Institute for Studies in Education

Pour ceux qui ont reçu le matériel - Parlons du Passé - nous aimerions savoir si, depuis décembre, vous avez pu ré-utiliser des unités du matériel dans votre enseignement, ou s'il a été possible de créer de nouvelles activités en vous servant des idées suggérées dans le matériel comme point de départ?

	Matériel ré-utilisé comme tel		Matériel utilisé comme point de départ		Nombre total d'heures
	oui	non	oui	non	
a) Le loup garou et le châle					
b) Passez-moi vos jumelles					
c) Vive la différence					
d) Proverbes					
e) Qui et Quoi					
f) Place à la fantaisie					
g) Casse-tête					
h) Souvenirs de mon enfance					
i) A la recherche					

Merci beaucoup de votre coopération.

L'Institut d'Etudes Pédagogiques de l'Ontario
The Ontario Institute for Studies in Education

Appendix D

Esquisse d'équivalences: français — anglais

Emplois	Français	Anglais
	Imparfait	simple past (-ed) past progressive (was, were + -ing used to + verb would + verb
	Exemples:	
(a) La description dans le passé	Luc était une homme bourru Il vivait près de la forêt L'Indienne aimait les animaux	Luke was a surly man He lived near the forest The Indian woman loved animals
(b) Une action habituelle ou qui se répète dans le passé	Il se mettait souvent en colère	often got He often used to get angry would often get
	Chaque matin elle se promenait dans la forêt	Every morning she went used to go for a walk in the forest would go
(c) Une action inachevée dans le passé	Son châle l'empêchait d'aller très vite	Her shawl prevented her from going very fast was preventing
	Elle marchait lentement, trop lentement	She was walking slowly, too slowly
(d) Une action en train de s'accomplir qui est interrompue	Tandis qu'ils déjeunaient, Luc a posé sa cuillère	As they were eating, Luke put down his spoon
(e) Un événement hypothétique	Si Luc me parlait plus souvent je ne me sentirais pas si isolée	If Luke spoke to me more often, I wouldn't feel so isolated
	Passé composé	simple past (-ed) present perfect (have, has + -en)
(f) Les principaux événements d'une narration	Marie-Rose a enlevé son châle, l'a rangé dans l'armoire et s'est mise au lit	Marie-Rose took off her shawl, put it away in the closet and got into bed
(g) Un événement du passé qui interrompt un autre événement en train de s'accomplir	Pendant qu'elle s'aventurait dans la forêt, elle a entendu le cri des loups-garous	While she was wandering in the forest, she heard the wolves howl
(h) Un événement qui a commencé à un moment défini du passé	Dès ce moment-là, les Indiens ont commencé à recueillir l'eau d'érable	From that moment on, the Indians began to collect maple syrup
(i) Une situation qui dure du passé jusqu'au présent	Il a toujours aimé la musique	He has always loved music

Appendix E

Le Passé en perspective

1) Cette légende raconte l'histoire d'un fermier qui s'appelait Luc. Il vivait près de la forêt avec sa femme Marie-Rose. Luc était un homme impatient et taciturne. Il n'aimait pas se mêler aux gens. Il se mettait souvent en colère.

2) Malgré tout cela, sa femme Marie-Rose gardait sa bonne humeur. Elle travaillait aux côtés de Luc sur la ferme et elle ne se plaignait jamais de son mari.

3) Un jour, Marie-Rose a decidé d'aller dans la forêt pour ramasser du petit bois.

4) Il faisait déjà un peu froid la nuit et le feu dans le foyer s'éteignait lentement. Luc n'était pas à la maison. Depuis quelque temps, il avait l'habitude de rentrer tard, ce qui inquiétait Marie-Rose.

5) Ce jour-là, Marie-Rose a pris son châle, l'a jeté sur ses épaules et a quitté la maison. Elle est entrée dans la forêt.

6a) Sans s'en apercevoir, elle s'aventurait plus avant dans la forêt

6b) quand, soudain, elle s'est rendu compte

6c) qu'il était très tard.

7) Elle a déposé un dernier morceau de bois dans son châle et s'est dirigée en toute hâte vers la maison.

8a) Beaucoup plus tard, tandis que Marie-Rose dormait à poings fermés,

8b) Luc est rentré et s'est mis au lit à son tour.

9a) Le lendemain matin tandis que Marie-Rose et Luc déjeunaient,

9b) Luc a soudain cessé de manger, a posé sa cuillère et a fait une grimace. Il a ouvert la bouche et Marie-Rose a retiré deux bouts de laine coincés entre ses dents - un rouge et un jaune. Tout a coup, elle a quitté la pièce et a couru vers l'armoire. Elle a ouvert l'armoire et a sorti son châle.

10) Marie-Rose n'en croyait pas ses yeux. Elle savait qu'elle tenait la vérité dans sa main.

11) Horrifiée, elle est retournée dans la cuisine voir son mari.

Appendix F

Compositions, scoring procedures and reliabilities

A

Nom: ______________________________

École: ______________________________

Classe: ______________________________

RÉDACTION FRANÇAISE

'Aux voleurs!'

On te demande d'écrire une rédaction d'une page.
Tu as 15 minutes pour le faire.

Aux voleurs!

L'été dernier, il y a eu un important vol de banque. Écris une histoire là-dessus, en utilisant les temps du passé. Commence avec le début suivant:

Ce jour-là, comme d'habitude, la banque était pleine de monde. Tout d'un coup, trois bandits ______________________________

B

Nom: ______________________

École: ______________________

Classe: ______________________

RÉDACTION FRANÇAISE

'Au secours!'

On te demande d'écrire une rédaction d'une page.
Tu as 15 minutes pour le faire.

Au secours!

L'été dernier, il y a eu un sauvetage d'un petit chat qui était pris dans un arbre. Écris une histoire là-dessus, en utilisant les temps du passé. Commence avec le début suivant:

C'était un beau dimanche d'été et sur le balcon de la maison des Dupont un petit chat dormait tranquillement. Tout d'un coup trois chiens ____________

Scoring of Compositions

Scoring focussed on the appropriate selection of the passé composé or imparfait according to context.

I. The following were defined as errors:

1. use of a tense other than the passé composé (or passé simple) in a context calling for the passé composé (or passé simple);
2. use of a tense other than the imparfait in a context for the imparfait;
3. use of an infinitive or second person plural (-ez) form instead of the passé composé or imparfait;
4. use of a mixed form (auxiliary with the imparfait, the passé simple, or an infinitive which is non-homophonous with its past participle) (see Part II, point 4, below).
5. use of the past participle without the auxiliary in a context for the passé composé or imparfait.

These errors were given equal weight in scoring.

II. The following were disregarded in scoring (i.e. not considered errors):

1. any obligatory context for a tense other than imparfait and passé composé (e.g., plus-que-parfait; présent in direct discourse);
2. in particular, use of imparfait or passé composé in a context requiring some other tense (e.g., présent, plus-que-parfait);
3. use of the passé simple instead of the passé composé;
4. spelling errors such as use of an -er infinitive or the 2nd person plural (-ez) form instead of the past participle (e.g., nous avons manger/mangez);
5. errors in the form of the past participle, especially but not exclusively with irregular verbs (e.g., buvé instead of bu; finé instead of fini);
6. errors in number or gender within a tense (e.g., nous allaient; la femme est allé);
7. addition, omission, substitution, or incorrect positioning of pronouns, including the reflexive pronoun;
8. spelling errors in the root of the verb;
9. use of the incorrect auxiliary (e.g., j'ai allé; nous sommes fait).

III. Scoring procedures:

Each composition was evaluated as described below, and assigned either a score from 1 to 5 or one of the three special codes (6, 7, 8).

1. Keeping in mind that the instructions set an overall context for use of past tenses, the scorers checked the composition for the number of obligatory contexts for the *imparfait* or *passé composé*, including those which occurred within direct discourse which was otherwise in the present.

   8 was assigned if there were fewer than 5 such obligatory contexts. The composition was not evaluated further.

2. If the student failed to use one of the *imparfait* and *passé composé*, but used the other with no errors in choice/form (as set out in Part I), a special code was assigned. (Obligatory contexts for tenses other than *imparfait* and *passé composé* were disregarded.)

   7 was assigned if the student used only the *imparfait*, and made no errors as defined in Part I.

   6 was assigned if the student used only the *passé composé*, and made no errors as defined in Part I.

3. Otherwise, the composition was scored on the following 5-point scale.

   The basis for scoring was the proportion of errors (as listed in Part I) in contexts for *imparfait* and *passé composé*.

   The following served as guidelines, not rigid definitions; it was expected that the scorer would use subjective judgement to estimate the level of error, rather than doing a precise count.

		Approximate Proport. errors
1	Errors occur in most or all obligatory contexts (performance is near chance, or worse)	(.75 - 1.)
2	Errors occur in many obligatory contexts (performance a little better than chance)	(.5 -.75)
3	Errors occur less frequently than correct choices/ forms (performance clearly better than chance)	(.25 - .5)
4	Errors occur much less frequently than correct choices/forms	(.25)
5	No errors, as defined in Part I.	

Reliability of Composition Measure

All compositions were scored by scorer F and most by scorer B, both of whom were native speakers of French. Inter-rater reliability was computed using compositions considered by both F and B to be scorable. It is reported for each form of the composition task at each phase of testing (pre-test, immediate post-test, and delayed post-test) in the accompanying Table F1. Since reliability was excellent for both forms and all three phases, and to minimize the number of cases lost due to missing data codes, cases of disagreement on scorability were rescored by scorer F. Only those compositions were then discarded for which two out of the three evaluations obtained indicated that they were 'unscorable'.

The final score for each composition was then computed as the mean of the two scorers' evaluations, or as scorer F's alone when none was available from scorer B. Unscorable compositions were treated as missing data.

Table Fl

Inter-rater Reliability on Composition Form A and Form B at each Phase of Testing

			Scorer F		Scorer B			
Phase	Form	N	Mean	sd	Mean	sd	F	
Pre-test	A	113	3.13	1.33	3.01	1.24	3.45	.92
	B	112	2.78	1.41	2.86	1.31	1.29	.92
Immediate Post-test	A	103	3.06	1.29	3.09	1.20	0.22	.93
	B	119	3.18	1.10	3.27	1.12	2.02	.91
Delayed Post-test	A	114	3.28	1.16	3.47	1.10	9.64**	.91
	B	117	3.14	1.25	3.26	1.29	4.38*	.94

* $p < .05$
** $p < .01$

Appendix G

A

Nom: ____________________

École: ____________________

Classe: ____________________

EXERCICE: 'LE MONSTRE'

Exemple:

Hier, pendant qu'il ____________________ dans le parc, l'orage
a couru / courait
____________________.
a éclaté / éclatait

© Ontario Institute for Studies in Education

LE MONSTRE

Hier soir, Charles ____________________ nous voir à la maison et il nous a raconté
est venu / venait
une histoire très bizarre. Veux-tu la connaître? Alors, la voici: Charles
______________________________ chez son oncle à la campagne. Son oncle
s'est reposé / se reposait
_________________________ de nombreux livres et tous les après-midis Charles
a possédé / possédait
____________________ parce qu'il aimait beaucoup les histoires surnaturelles. Un jour,
a lu / lisait
alors qu'il ____________________ assis dans le salon lisant comme d'habitude une histoire
a été / était
surnaturelle, il __________________ une expérience extraordinaire.
a eu / avait

Charles était encore plongé dans son histoire lorsque, distraitement, il
__________________ les yeux vers la fenêtre ouverte. Soudain, devant lui, il
a levé / levait
_________________ un monstre qui _________________________ peu à peu la colline.
a vu / voyait *a descendu / descendait*
Tout de suite Charles a pensé qu'il _________________________ fou. Pourtant il avait vu
est devenu / devenait
le monstre, il __________________ même le décrire: une tête énorme, une bouche à
a pu / pouvait
l'extrémité d'une longue trompe avec de chaque côté des défenses brillantes. Mais le
plus étrange, _____________________ la tête de mort dessinée sur sa poitrine. Au
ça a été / c'était
moment où Charles __________________ les mâchoires s'ouvrir, il __________________
a vu / voyait *a crié / criait*
de peur. Son oncle ________________________ à toute vitesse. Tout surpris, il
est arrivé / arrivait
______________________ à Charles: "Que se passe-t-il ici?" Après avoir écouté les
a demandé / demandait

explications de Charles, il ____________________: "Tu lis trop, tu devrais sortir et
[✓ a ajouté] / ajoutait
faire de l'exercice. Toute cette histoire n'est qu'un rêve".

Mais Charles ________________ sûr d'avoir vu le monstre et depuis cette vision,
a été / [✓ était]
il avait vraiment peur. Pour lui, ____________________ le signe que quelque chose de
ça a été / [✓ c'était]
terrible arriverait.

Quelques jours plus tard, alors que Charles et son oncle
________________________________ dans le même salon, Charles
se sont trouvés / [✓ se trouvaient]
________________ le monstre pour la seconde fois. En même temps, il
[✓ a vu] / voyait
______________________ le bras de son oncle et lui ____________________ la
[✓ a saisi] / saisissait *[✓ a montré] / montrait*
fenêtre. Son oncle _______________________ mais il n'y _______________ déjà plus
[✓ a regardé] / regardait *a eu / [✓ avait]*
rien. Tout de suite son oncle l' ___________________________ sur la forme, la
[✓ a interrogé] / interrogeait
couleur et la grosseur du monstre. Comme Charles ______________________, l'oncle
a répondu / [✓ répondait]
a souri. Il ___________________ prendre un livre dans la bibliothèque et il
[✓ est allé] / allait
_______________ à haute voix: "Sphinx, famille des Crépusculaires, classe des
[✓ a lu] / lisait
Insectes, bouche en forme de trompe, mandibules en forme de défense, sur le corps
dessin d'une tête de mort".

Cela décrivait exactement le monstre. Après un moment, il
______________________ le livre et il ___________________ s'asseoir à la place de
[✓ a refermé] / refermait *[✓ est venu] / venait*

Charles. Pendant qu'il ______________________ par la fenêtre il s'est exclamé: "Ah,
a regardé / [✓ *regardait*]

le voilà! Il est sur la colline. Oh, un bel animal, mais pas aussi gros que tu

__________________. Il avance maintenant sur un fil d'araignée le long de la fenêtre.
as pensé / [✓ *pensais*]

Ton oeil _________________________, la bête était trop près de toi, et le livre que tu
[✓ *t'a trompé*] / *te trompait*

__________________ trop impressionnant. Il faut se méfier des illusions d'optique
as lu / [✓ *lisais*]

Charles!"

Charles _______________________ son histoire avec un sourire mystérieux et il
[✓ *a terminé*] / *terminait*

nous _______________________ si nous avions déjà eu une telle expérience. Et toi,
[✓ *a demandé*] / *demandait*

t'est-il déjà arrivé quelque chose de pareil?

B

Nom: ____________________

École: ____________________

Classe: ____________________

EXERCICE: 'LE VOL'

Exemple:

Hier, pendant qu'il ____________________ dans le parc, l'orage
a couru / courait

____________________.
a éclaté / éclatait

© Ontario Institute for Studies in Education

LE VOL

Ce matin-là, en s'approchant des bâtiments de la Sécurité Nationale monsieur Leblanc a remarqué quelque chose d'anormal devant le laboratoire de physique. Il y

____________________ des voitures noires rangées côte à côte et tout le monde
a eu / avait

____________________ de l'événement: " Vers onze heures hier soir, le gardien de nuit a
a parlé / parlait

été attaqué par des inconnus. On ____________________ le pauvre homme
a retrouvé / retrouvait

inconscient dans le corridor du département de physique. On l' ____________________
a emmené / emmenait

d'urgence à l'hôpital."

Monsieur Leblanc qui ____________________ avoir plus de détails sur l'affaire,
a voulu / voulait

____________________ dans le bâtiment. Au même instant, son ami le
s'est précipité / se précipitait

professeur Pelletier ____________________ en courant, suivi de plusieurs personnes.
est arrivé / arrivait

"On ____________________ mes dossiers secrets, le résultat de 15 ans de recherche!"
a volé / volait

a crié le professeur, qui ____________________ pour la Sécurité Nationale. Il
a travaillé / travaillait

venait d'inventer un remède efficace contre les radiations atomiques.

À ce moment, un journaliste assistant à la scène ____________________ l'interroger:
est venu / venait

"Avez-vous des soupçons?" a-t-il demandé. "Non" ____________________ le
a répondu / répondait

professeur, "Mais les voleurs ____________________ une clé du laboratoire
ont possédé / possédaient

et ils ___(item deleted)___ savoir exactement où ____________ mes
ont dû / devaient *ont été / étaient*

papiers, car ils n'____________ qu'un tiroir.
ont ouvert / ouvraient

- Qui ____________ au courant?" ____________ le reporter.
a été / etait *a demandé / demandait*

- Personne, absolument personne." Soudain le visage du professeur

____________ de couleur et il ____________
a changé / changeait *s'est retourné / se retournait*

brusquement vers son confrère, monsieur Leblanc, qui, lui,

____________.
a sursauté / sursautait

Plus tard en parlant avec sa femme, le professeur Pelletier a dit: "Quand la police m'a interrogé, je n'ai pas voulu leur faire part de mes soupçons. Claude, mon jeune

assistant, qui avait toute ma confiance, ____________ l'importance de mes
a connu / connaissait

travaux. Il ____________ la seule personne à savoir où se
a été / était

____________ mes papiers. Le matin du vol, il
sont trouvés / trouvaient

____________ dans mon bureau, au moment où André Leblanc et moi, nous
est entré / entrait

consultions mes notes. Nous ____________ rendez-vous tous les trois avec le
avons eu / avions

Directeur de la Sécurité Nationale, alors j' ____________ mes dossiers, je
ai arrangé / arrangeais

les ____________ dans le tiroir devant lui, et puis nous
ai mis / mettais

____________." Ainsi tout ____________ Claude comme le
sommes sortis / sortions *a désigné / désignait*

coupable.

Après plusieurs semaines d'enquête, l'inspecteur de police

______________ le professeur Pelletier. "Voilà votre coupable" a annoncé
[✓ a convoqué] / convoquait

l'inspecteur, en désignant un homme assis au fond du bureau. "On

______________ ses empreintes à l'intérieur du tiroir dont vous seul
[✓ a relevé] / relevait

______________ la clé. On ______________ l'outil qu'il
avez possédé / [✓ possédiez] *[✓ a retrouvé] / retrouvait*

______________ pour forcer la serrure et on ______________
[✓ a utilisé] / utilisait *[✓ a découvert] / découvrait*

tous vos dossiers dans sa valise, alors qu'il les ______________ à ses
a apportés / [✓ apportait]

complices contre une forte somme d'argent."

Le professeur, ______________ la pièce rapidement, accablé par les
[✓ a quitté] / quittait

preuves que l'inspecteur lui donnait. Il ne pouvait toujours pas y croire: André Leblanc,

un des ses plus anciens collègues et amis, ______________ partie d'un réseau
a fait / [✓ faisait]

d'espionage!

Reliability of Cloze Measure

Inter-item reliability was computed on the pre-test administration of Form A and Form B separately. Results are shown in the accompanying Table G1. The sample for each form includes only those students who completed the cloze test in all three phases in their assigned order.

Table G1

Inter-item Reliability of Cloze Form A and Form B at Pre-test.
(Based on Students who completed the cloze test for all three phases of testing.)

Test	No. of Cases	No. of Items	F	
Form A	137	37	19.22**	.610
Form B	133	37	26.57**	.685

**$p < .01$

Appendix H

ENTREVUE - FORME A

Les préliminaires

1. **Comment t'appelles-tu?**
2. **Assieds-toi ici, etc.**

L'entrevue propre

1. **Qu'est-ce que ta classe faisait au moment où tu es parti pour venir ici?**

2. **Peux-tu me raconter quelque chose d'amusant qui s'est passé à l'école depuis la rentrée? Quand tu étais en cinquième année peut-être?**

   Si l'élève ne se rappelle pas, ajoutez:

   Peux-tu me raconter quelque chose d'amusant qui est arrivé chez toi, ou avec tes amis?

3. **Aimes-tu regarder la télé? Quelle est ton émission préférée?**

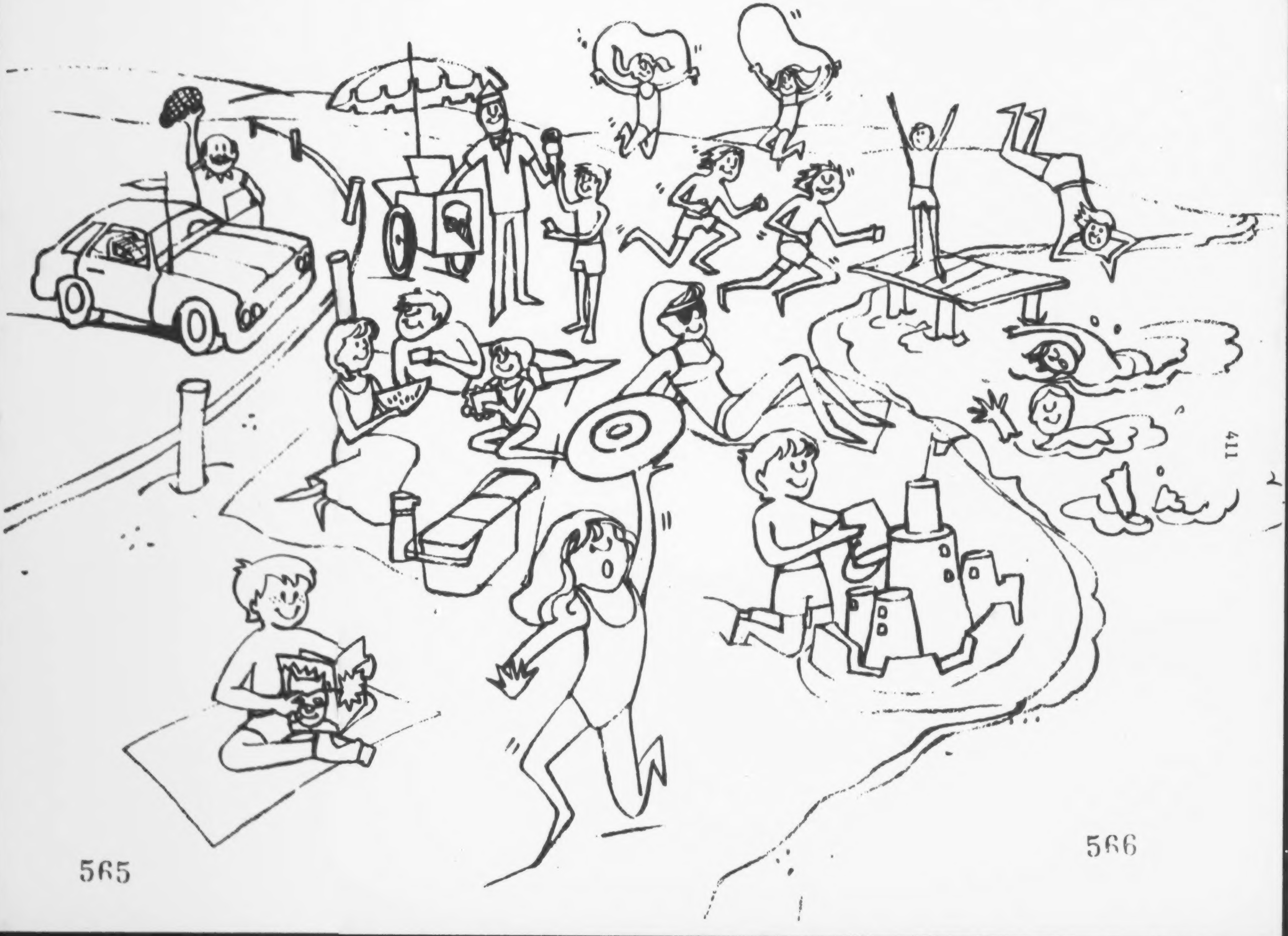
4. **Que penses-tu que les enfants faisaient autrefois quand il n'y avait pas de télé?**

5. **Voici un dessin qui représente un après-midi à la plage l'été dernier.** (Montrez le dessin de la plage.)

   Il faisait beau ce jour-là. Alors tout le monde s'amusait. Décris ce qui se passait quand l'oncle Marcel est arrivé en voiture. Bon, on commence avec Christine (avec le frisbee). **Qu'est-ce qu'elle faisait cet après-midi là quand l'oncle Marcel est arrivé?** Si l'élève se sert du présent, dites: **N'oublie pas qu'on ne parle pas de maintenant, on parle de l'été dernier.**

   Si l'élève demande une traduction, donnez-lui l'item en disant: **le verbe, c'est xxx.**

   Quand l'élève a fini de parler, vous dîtes: **C'est tout?**



565

566

ENTREVUE - FORME B

Les préliminaires

1. **Comment t'appelles-tu?**
2. **Assieds-toi ici, etc.**

L'entrevue propre

1. **Qu'est-ce que ta classe faisait au moment où tu es parti pour venir ici?**

2. **Es-tu jamais allé(e) au cirque? Au Ex? (C.N.E.) A la Ronde? Raconte-moi ce qui s'est passé.**

   Si l'élève n'a pas fait de telle visite, demandez: **Qu'est-ce que tu as fait pour ton anniversaire?**

3. **Vas-tu quelquefois avec ta mère ou ton père faire les courses au supermarché? Quel est ton repas préféré?**

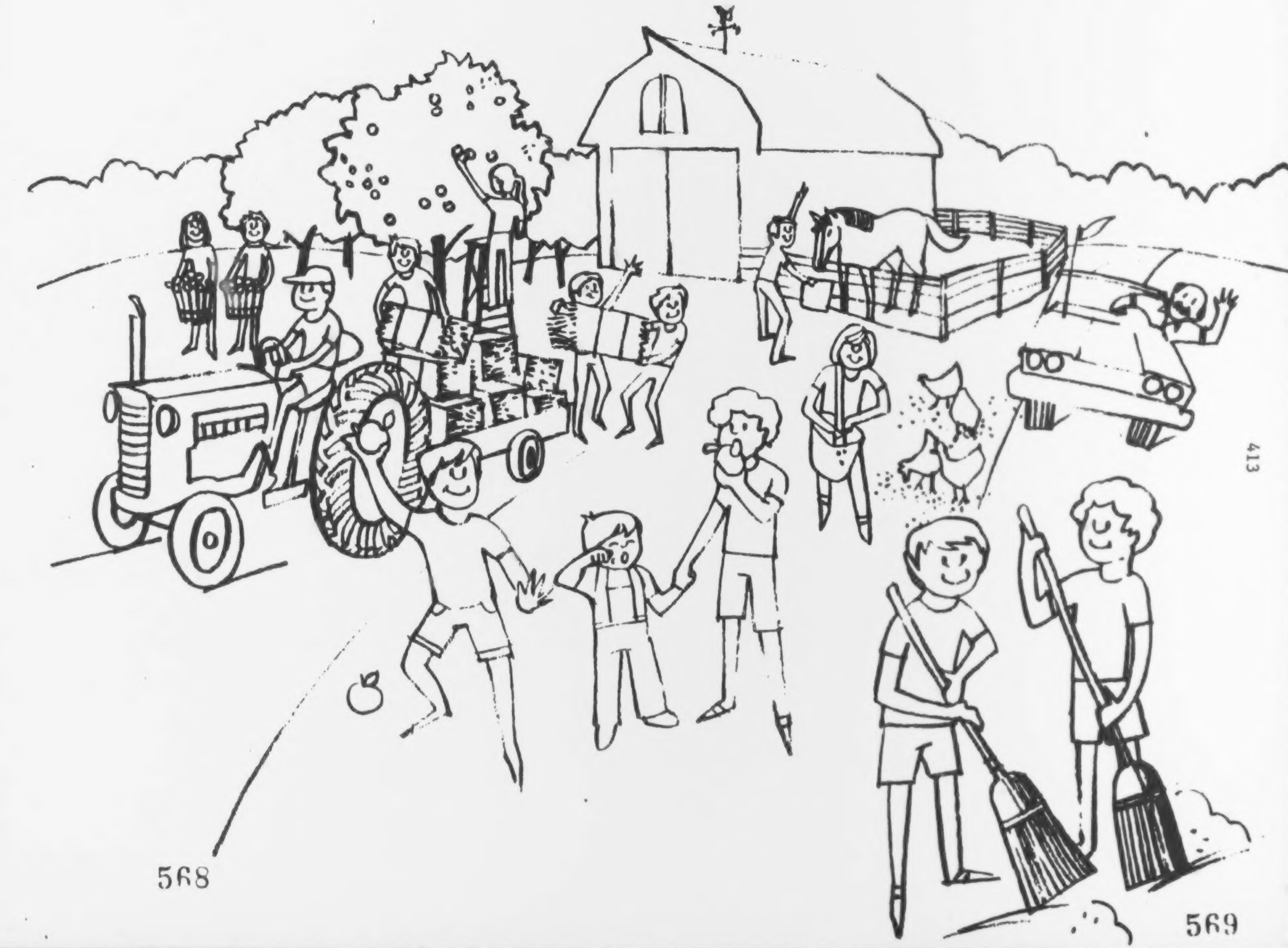
4. **Autrefois, au temps des pionniers, quand il n'y avait pas de supermarchés, qu'est-ce que tu penses qu'on faisait pour se nourrir? - Et quoi d'autre?**

5. **Voici un dessin qui représente un après-midi à la ferme l'été dernier.** (Montrez le dessin de la ferme.) **Il faisait beau ce jour-là. Alors tout le monde s'amusait. Décris ce qui se passait quand l'oncle Marcel est arrivé en voiture. Bon, on commence avec Pascal** (l'enfant pleurant). **Qu'est-ce qu'il faisait cet après-midi là quand l'oncle Marcel est arrivé?**

   Si l'élève se sert du présent, dites: **N'oublie pas qu'on ne parle pas de maintenant, on parle de l'été dernier.**

   Si l'élève demande une traduction, donnez-lui l'item en disant: **le verbe, c'est xxx.**

   Quand l'élève a fini de parler, dîtes: **C'est tout?**


568
569
413

Oral Interviews: Transcription, Scoring and Inter-rater Reliability

Relevant parts of students' responses to Questions 1, 4, and 5 of each interview were first transcribed in regular French orthography by two near-native speakers of French (M and S), working independently. Casual conversation related to the questions, hesitations, and strings of phrases containing no verbs were not transcribed. Particular attention was given to verb forms. Where there was ambiguity as to whether a student had used, e.g., an -er infinitive form or an imparfait form, the following criteria were used:

(a) If an overt grammatical subject was provided by either the student or the interviewer, the associated verb form was transcribed as the imparfait;

(b) If no subject was provided, the verb form was transcribed as an infinitive.

Coding was used to indicate the following:

(a) obligatory contexts for the imparfait (established with reference to the interview question and internal constraints in student responses);

(b) errors in these contexts (use of an infinitive or tensed form other than the imparfait);

(c) repetitions of verb forms supplied by the interviewer;

(d) self-corrections by the students; and

(e) requests for clarification.

Interviewer speech was not transcribed except to indicate prompts. Based on the transcripts, the number of obligatory contexts for the imparfait and the number of errors produced in these obligatory contexts were then counted for each of Questions 1, 4, and 5. Not included in these counts were:

(a) student talk following an interviewer prompt (asking for more information) in relation to Question 1;

(b) contexts and verb forms occurring in requests for clarification;

(c) use of 'il y a ' or 'il y avait' or the expression 'Il garde comme';

(d) repetitions by the student of verb forms supplied by the interviewer;

(e) earlier versions of forms that were then self-corrected by the student (i.e. only the final version was counted).

After several initial scoring sessions in which disagreements between the scorers were discussed, inter-rater reliabilities were calculated for 15 compositions (see Table H1). It was noted that the few remaining disagreements were almost entirely due to discrepancies in the initial transcription. Scoring procedures were therefore revised: Half the students in each school were assigned to each of the two scorers. Each tape was transcribed by the other scorer or by a native French speaker; the assigned scorer then checked and scored the transcript. All of the pre-test data were scored, then the immediate post-test, and finally the delayed post-test. Scorers consulted in case of disagreement on the transcript, but no further check on reliability was done.

For analyses, an error rate was computed from the number of errors on Questions 1, 4 and 5, divided by the number of obligatory contexts on these questions.

Table H1

Inter-rater Reliability of Oral Measure (N. of Cases = 15)

	Difference between Scorers					
	Scorer M		Scorer S			Correlations between
Variable	Mean	sd	Mean	sd	F	Scorers
Q.1 Errors	0.47	0.64	0.60	0.74	2.15	.88**
Q.1 Obligatory Contexts	1.53	0.83	1.67	0.98	2.15	.94**
Q.4 Errors	2.13	1.46	2.27	1.75	1.00	.96**
Q.4 Obligatory Contexts	3.47	0.99	3.47	1.13	0.00	.88**
Q.5 Errors	7.87	3.11	7.67	3.13	0.32	.90**
Q.5 Obligatory Contexts	12.47	2.62	12.47	2.90	0.00	.97**
Error Rate, Q1,4,5	0.60	0.25	0.60	0.24	0.00	.94**

**p<.01

THE DEVELOPMENT OF BILINGUAL PROFICIENCY

FINAL REPORT

Volume III: Social Context and Age

Birgit Harley
Patrick Allen
Jim Cummins
Merrill Swain

April 1987

Modern Language Centre
The Ontario Institute for Studies in Education

TABLE OF CONTENTS

Page

VOLUME I

Preface vii
Project Staff viii

Introduction
THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. Purpose 1
2. The Nature of Language Proficiency 1
3. Classroom Treatment Studies 8
4. Social Context and Age 13
5. Conclusions 19

Chapter 1
PERSPECTIVES ON LEXICAL PROFICIENCY IN A SECOND LANGUAGE

1. Introduction 23
2. The Data Base 24
3. Measures of Lexical Proficiency 26
4. Lexical Measures and L2 Proficiency 30
5. Lexical Use of Immersion Students and Native Speakers 34
6. Conclusion 40

Tables 44
Figures 60
Appendix A 64
Appendix B 66

Chapter 2
THE DEVELOPMENTAL GROWTH OF METAPHOR COMPREHENSION IN CHILDREN'S FIRST AND SECOND LANGUAGE

1. The Development of Bilingualism 67
2. A Cognitive Developmental Perspective 69
3. Summary of the Present Study and Hypotheses 70
4. Method 71
5. Results: Metaphors Interpreted in English 76
6. Results: Metaphor Interpretations in Spanish 81
7. Conclusions 84

Tables 91
Figures 101
Appendix 107

Page

VOLUME II

Preface vii
Project Staff viii

Introduction
THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. Purpose 1
2. The Nature of Language Proficiency 1
3. Classroom Treatment Studies 8
4. Social Context and Age 13
5. Conclusions 19

Chapter 3
COLT OBSERVATION SCHEME: DEVELOPMENT AND VALIDATION

1. Introduction 23
2. Description and Rationale of the Observation Scheme 24
3. The Validation Study 36
4. Discussion 43

Tables 50
Figures 52
Appendix A 54

Chapter 4
THE CORE FRENCH OBSERVATION STUDY

1. Introduction 56
2. The Sample 57
3. General Procedures 59
4. The Colt Observation Scheme: Procedures 61
5. Classroom Practice: Findings 64
6. Proficiency Predictions 70
7. Proficiency Tests 72
8. Process/Product Findings 78
9. Discussion and Interpretation of Results 86

Tables 98
Appendix A 127
Appendix B 137
Appendix C 161
Appendix D 164
Appendix E 167
Appendix F 181

Page

VOLUME II (continued)

Chapter 5
THE IMMERSION OBSERVATION STUDY

1. Introduction 190
2. Vocabulary Instruction in Immersion Classes 192
3. Vous/Tu Input 222
4. Student Talk in Teacher-Fronted Activities 226
5. Error Treatment 233
6. Concluding Comment 240

Tables 252
Appendix A 264
Appendix B 287
Appendix C 293
Appendix D 317

Chapter 6
FUNCTIONAL GRAMMAR IN FRENCH IMMERSION:
A CLASSROOM EXPERIMENT

1. Purpose and Rationale 342
2. Design of the Experiment 346
3. The Findings 354
4. Discussion 361
5. Conclusion 363

Tables 367
Figures 379
Appendix A 381
Appendix B 385
Appendix C 386
Appendix D 389
Appendix E 390
Appendix F 391
Appendix G 399
Appendix H 409

Page

VOLUME III

Preface vii

Project Staff viii

Introduction

THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. Purpose 1
2. The Nature of Language Proficiency 1
3. Classroom Treatment Studies 8
4. Social Context and Age 13
5. Conclusions 19

Chapter 7

THE LANGUAGE USE PATTERNS, LANGUAGE ATTITUDES, AND BILINGUAL PROFICIENCY OF PORTUGUESE CANADIAN CHILDREN IN TORONTO

1. Introduction 23
2. Scoring Procedures and Refinement of the Data Set 28
3. Results 35
4. Comparison of Portuguese Proficiency with Azorean Native Speakers 49

Tables 54

Appendix A 85

Appendix B 88

Appendix C 106

Appendix D 119

Appendix E 154

Chapter 8

THE DEVELOPMENT OF BILINGUAL PROFICIENCY IN THE TRANSITION FROM HOME TO SCHOOL: A LONGITUDINAL STUDY OF PORTUGUESE-SPEAKING CHILDREN

1. Purpose of the Study 159
2. Methodology 159
3. Scoring Procedures and Preliminary Results for Home Interaction Data 165
4. Home Interaction of 4-Year Old Portuguese Background Children in Toronto: Preliminary Trends 176

Tables 182

Page

VOLUME III (continued)

Chapter 9
AGE OF ARRIVAL, LENGTH OF RESIDENCE, AND INTERDEPENDENCE OF LITERACY SKILLS AMONG JAPANESE IMMIGRANT STUDENTS

1. Background 183
2. Methodology 185
3. Results 188
4. Conclusion 191

Tables 194
Figures 199
Appendix A 200

Chapter 10
THE RELATIONSHIP BETWEEN STARTING AGE AND ORAL SECOND LANGUAGE PROFICIENCY IN THREE GROUPS OF CLASSROOM LEARNERS

1. Introduction 203
2. The Sample 204
3. Procedures 205
4. Verb Use: Analysis and Results 207
5. Fluency: Analysis and Results 213
6. Sociolinguistic Test Results 216
7. Discussion 216
8. Conclusion 220

Tables 223
Appendix A 232
Appendix B 235
Appendix C 253
Appendix D 251

PREFACE

The Development of Bilingual Proficiency is a large-scale, five-year research project which began in September 1981. The present Final Report of the project is the third in a series. It follows an interim Year 1 Report, produced in September 1982 at the end of the first year of the project, and a Year 2 Report, produced in September 1983.

There are three volumes in this Final Report of the project, each concentrating on specific issues investigated in the research: the nature of language proficiency (Volume I), the effect of classroom treatment on language proficiency (Volume II), and the relevance of social context and age for language learning (Volume III). Each volume is introduced by an identical 20-page overview of all the studies carried out in the context of the Development of Bilingual Proficiency (DBP) Project. The overview includes brief summaries of the individual studies together with an indication as to where the complete report of each study is to be found (either in the Year 2 Report or in Volume I, II, or III of the Final Report). Within the complete reports of individual studies contained in this Final Report, references to other Development of Bilingual Proficiency Project studies appear either as 'Year 2 Report' or, when they form part of the Final Report, as chapter numbers only. Note that Chapters 1 and 2 appear in Volume I, Chapters 3-6 in Volume II, and Chapters 7-10 in Volume III.

The authors wish to acknowledge the contribution of the many individuals and organizations who have played a role in the Development of Bilingual Proficiency Project since it began in September 1981. We are grateful to administrators, teachers, students and their parents in the following Ontario school boards for their participation in the research: the Board of Education for the City of Scarborough, the Carleton Board of Education, the Metropolitan Separate School Board, the North York Board of Education, and the Toronto Board of Education. In addition, we would like to thank the Portuguese Secretary of State for Immigration, the Regional Secretary of Social Affairs for the Autonomous Region of the Azores, and the staff, parents and students of the Japanese School of Toronto Shokokai Inc. We wish to acknowledge the contribution of Ellen Bialystok and Raymond Mougeon, who were principal investigators of the project in 1981-82 and 1981-83 respectively. We would also like to express our appreciation to project staff for their part in carrying out the research and in text-processing. Finally, we wish to acknowledge the financial support provided in the form of a five-year negotiated grant (No. 431-79-0003) by the Social Sciences and Humanities Research Council of Canada, and the administrative and financial contribution of the Ontario Institute for Studies in Education.

B.H., P.A., J.C., M.S.

PROJECT STAFF

(from September 1983)*

Principal Investigators:

Patrick Allen, Jim Cummins, Birgit Harley, Merrill Swain

Project Managers:

Birgit Harley, Sharon Lapkin

Research Officers:

Margarida Aguiar, Jud Burtis, Susanne Carroll, Maria Fröhlich, Vince Gaudino, Mary Lou King, Laurette Lévy, José Lopes, Françoise Pelletier, Bruno Piché, Jorge Ramos, Mireille Tremblay, Virginia Won (Pang)

Graduate Assistants:

Jane Appelt, Eva Bild, Patrick Conteh, Giselle Corbeil, Vera Held, Richard Kruk, Bill Marshall, Josée Mazzilli, Anne Riddle, Emilia Rivas-Arraiz, Kathryn Shred, Nina Spada, Yilin Sun, Tim Tomlinson

Post-doctoral Fellows:

Monica Heller, Janice Johnson

Statistical Consultant:

Jamshid Etezadi-Amoli

Text-processing Specialist:

Ellen Jeske

Secretaries:

Joyce Howlett, Doreen Panowyk, Donna Schillaci

---

* For staff members in previous years, see earlier reports

Introduction

THE DEVELOPMENT OF BILINGUAL PROFICIENCY PROJECT: OVERVIEW OF STUDIES

1. PURPOSE

The purpose of this five-year research project has been to investigate issues concerning language proficiency[1] and its development in educational contexts for children learning a second language. The research has concentrated on the following major issues: the nature of language proficiency; the impact of instructional practices on language learning; the relationship between social-environmental factors and bilingual proficiency; and the relationship between age and language proficiency. In this overview of the project, studies focussing on each of these issues are summarized.

2. THE NATURE OF LANGUAGE PROFICIENCY

The focus and ultimate goal of all studies carried out within the Development of Bilingual Proficiency Project is the improvement of educational practices as they relate to second language learning and teaching. Because so much of school practice relates rather narrowly to the teaching and learning of grammatical aspects of the target language, it was considered essential to broaden the scope of the typical educational definition of language proficiency to incorporate discourse and sociolinguistic dimensions, and to consider the differential demands that context-reduced versus more context-embedded language tasks may make on the learner.

2:1 Large-scale Proficiency Study
(Year 2 Report)

Objectives. The primary purpose of the large-scale proficiency study conducted during the first two years of the Project was to determine whether the three hypothesized traits, representing key components of language proficiency, could be empirically distinguished. It was hypothesized that grammatical, discourse, and sociolinguistic competence would emerge as distinct components of second language proficiency which may be differentially manifested under different task conditions. A secondary purpose of the study was to develop a set of exemplary test items and scoring procedures that could be used, or modified for use, in further studies involving the measurement of the hypothesized traits. A final purpose of the study was to provide a broadly based description of the target language proficiency of the second language learners tested, in relation to that of native speakers.

Subjects. A total of 198 students was involved in the study. Of these, 175 were grade 6 early French immersion students from the Ottawa region, and 23 were grade 6 native speakers from a regular Francophone school in Montreal. The immersion students, in six intact classes, had received 100% of their schooling in French in kindergarten to grade 2 or 3, since when they had been taught in English for a gradually increasing portion of each day. At the time of testing, about 50% of their school subjects were

being taught in French, and the other 50% in English. This sample of classroom second language learners was selected because of the theoretically interesting and educationally innovative nature of their intensive school-based language learning experience, and because they were at an age where they were sufficiently proficient in the second language to be able to cope with a wide range of types of language tasks.

Instruments. A multi-method multi-trait design was used to determine the extent to which grammatical, discourse, and sociolinguistic dimensions of the immersion students' French proficiency were distinguishable. To measure proficiency on each trait, three methods of testing were used: oral production, multiple choice, and written production. A matrix with nine test cells was thus created, consisting of three tests of grammar, three of discourse, and three of sociolinguistics. The oral production task for each trait was administered to a randomly selected sub-sample of 69 immersion students and ten native speakers, representing ten-eleven subjects from each class.

Grammatical competence was operationalized for the purposes of this study as rules of morphology and syntax, with a major emphasis on verbs and prepositions. The grammar oral production task consisted of a guided individual interview in which the interviewers' questions were designed to elicit a variety of verb forms and prepositions in French, as well as responses that were sufficiently elaborated to score for syntactic accuracy. The content of the interview questions (e.g. favourite pastimes, trips taken) was at the same time designed to focus the subject's attention on communication rather than the code. Grammatical scoring was based on the student's ability to use certain grammatical forms accurately in the context of particular questions. The group-administered grammar multiple choice test consisted of 45 written items which also assessed knowledge of the verb system, prepositions, and other syntactic rules, including written agreement rules. The student's task was to select the correct response from three alternatives provided. The third grammar task, written production, consisted of two short compositions to be written in 15 minutes each — one a narrative and the other a letter of request. Both this written production task and a parallel discourse written production task — also involving a narrative and a request letter — were assessed for grammatical proficiency. Scoring focussed on grammatical accuracy in verbs, prepositions, and other rules of syntax and morphology.

The discourse trait was defined as the ability to produce and recognize coherent and cohesive text (written or oral). For the individual discourse oral production task, the student was required to retell the story of a silent movie and to present arguments in support of an opinion. This task was rated on 5-point scales both globally and in detail for coherence and cohesion, focussing, for example, on the student's ability to make clear and accurate reference to characters, objects, and locations, to produce a logically connected text, and to fulfill the basic task required. The discourse multiple choice task consisted of 29 short written passages from each of which a sentence had been omitted. The student was required to select from three alternatives the sentence that best fit the context. The discourse written production task, like the grammar written production task, consisted of a narrative and a request letter. All four (grammar and discourse) tasks were rated for proficiency in discourse on the same kinds of features that were assessed in the discourse oral production task.

Sociolinguistic competence, the third trait dealt with in this proficiency study, was operationalized as the ability to produce and recognize socially appropriate language in context. The individual oral production task involved a set of slides with taped descriptions representing situations of different levels of formality. The student's task

was to respond appropriately with a request, offer, or complaint. Scoring focussed on the student's ability to shift register appropriately. Thus sociolinguistic proficiency was measured by difference scores, calculated by subtracting the number of formal 'politeness' markers produced by the student in informal variants of situations from those produced in formal variants of the situations. The sociolinguistic multiple choice test consisted of 28 items, each with three alternative ways of expressing a given sociocultural function. The choices were all grammatically accurate but not equally appropriate. The student's task was to select the most appropriate of the choices in the given situation. Scoring of responses was weighted according to the choices made by native speakers. The sociolinguistic written production task involved the writing of a formal request letter and two informal notes, all of which could be categorized as directives. The request letter written as part of the discourse written production task was also scored for sociolinguistic proficiency. As for the oral production task, scoring was based on difference scores, calculated by subtracting the number of formal markers produced in the notes from those produced in the letters.

Reliability and generalizability of scores. The component within-test scores were combined to produce a single overall score for each of the nine trait-method cells in the matrix. The composition of each of these overall scores was calculated to maximize validity and reliability. On the multiple choice tests, the reliability of the immersion students' total scores ranged from .58 on the sociolinguistic test to .75 on the discourse test. Generalizability studies were conducted on those cells for which sufficient data were available: the sociolinguistic oral production test and the three written production tests. G-coefficients for these tests, based on the subsample of orally tested students, were comparable to the multiple choice test reliabilities.

Testing a model of proficiency. In order to determine whether the three traits -- grammatical, discourse, and sociolinguistic competence -- could be empirically distinguished, two kinds of analyses were performed: (a) factor analysis, and (b) a comparison of the group means of the learners and native speakers.

The factor analysis based on the 69 orally tested immersion students failed to confirm the hypothesized three-trait structure of proficiency. Instead, confirmatory factor analysis by means of LISREL produced a two-factor solution. One of these factors was interpretable as a general language proficiency factor; it had positive loadings from all cells in the nine-test matrix except for the sociolinguistic written production test. The highest loadings on this general factor were from the three grammatical tests. The second factor was interpretable as a written method factor; it had loadings from the three multiple choice tests and from all three written production tests. The tests loading on this method factor appeared to be tapping the kind of literacy-oriented linguistic proficiency that is typically learned in classrooms. The lack of trait structure emerging from the factor analysis may have been due to the homogeneity of the immersion student sample. Within their classroom setting these students had all had very much the same kind of exposure to French, and strong opportunities for some students to develop proficiency in one area, and other students to develop proficiency in a different area, were lacking.

A different kind of result emerged from comparisons of immersion and native-speaker scores on the various tests. On all three grammar tests, the immersion students' mean score was considerably lower than that of the native speakers (p .01), and they also scored generally lower on the sociolinguistic tests than did the native speakers. On the discourse tasks, however, the scores of the immersion students were close or

equivalent to those of the native speakers and there were no significant between-group differences. The immersion students' strong performance in discourse may have been due to positive transfer from prior experience in their mother tongue. In contrast to the factor analysis results, then, these comparative findings, showing very different results for discourse as opposed to grammar and sociolinguistics, provide some evidence in support of a distinction between traits.

Conclusions. It was concluded that, although the three hypothesized language proficiency traits were not empirically distinguished via the factor analysis, this result may have been dependent on the relatively homogeneous language learning background of the immersion population studied. This did not necessarily mean that the traits would not be distinguishable in a more heterogeneous language learning population. From an educational perspective it was clear that the analysis of proficiency into different components was diagnostically revealing of the second language strengths and weaknesses of the immersion students. It was decided that two kinds of further studies were indicated to probe issues concerning how different dimensions of proficiency develop as a function of the immersion students' specific language learning experience: (a) small-scale in-depth studies of specific aspects of the immersion students' second language proficiency based on the data already collected (see 2:2 - 2:3 below), and (b) the study of language learning activities in the immersion classroom setting (see 3:3 - 3:4 below).

2:2 Transfer in Immersion Students' Compositions (Year 2 Report)

Hypotheses and design. Given the shared mother tongue, English, of the immersion students and the dominance of English in the wider school and outside-school environment of the immersion program, mother tongue transfer was expected to be a continuing factor in the students' written production at the grade 6 level. In a small-scale study of compositions written by 22 native speakers and 22 of the orally tested immersion students from two randomly selected classes in the larger proficiency study, evidence was sought for the hypothesis that mother tongue transfer may be manifested in the way in which the learners were distributing semantic information across syntactic elements in the second language, without necessarily making outright errors.

One of the composition topics assigned in the large-scale proficiency study, Au secours!, involved writing a story about the rescue of a kitten from a tree. The students' stories on this topic contained a very similar series of events, involving several changes of location. The focus of the present study was on how the immersion students were expressing the location/direction distinction in these stories, given that there are characteristic differences between French and English in this linguistic domain. While in English, prepositions generally serve an important role in conveying the location/direction distinction (e.g. at/to, in/into), in French there is a general tendency for direction to be expressed in the verb, and for prepositions (e.g. à, dans, sur) to be neutral with respect to the location/direction distinction. It was hypothesized that the immersion students would rely on prepositions rather than the verb to express the notion of direction.

Findings. A comparison of directional expressions in the Au secours! stories written by the immersion students and the native speakers showed that, as expected, the immersion students were much less likely than the native speakers to mark direction in the verb, preferring a non-directional verb of motion such as courir together with a

preposition to express the directional notion. The immersion students, at the same time, sometimes erroneously used French prepositions unmarked for direction as if they were carrying the directional distinction, and also tended to make more use than the native speakers of prepositional phrases expressing direction, even on those occasions when they also used directional verbs. This latter tendency did not necessarily lead to error. The findings of the study thus provide support for the hypothesis that the immersion students would show a systematic tendency to rely more heavily on prepositions to express the notion of direction than the native speakers.

Conclusions. It was concluded that the students may need more focussed classroom input that would alert them to such characteristics of French that are different from English, together with more opportunities for expressing the relevant distinctions in their second language.

2:3 Lexical Proficiency in a Second Language (Final Report, Vol. I)

In the large-scale proficiency study described above (2:1), there were no measures specifically designed to assess lexical proficiency, not because lexical proficiency was considered unimportant but because it was assumed to enter into performance on all the tasks assigned. In the present study, the two narratives and three request letters written by 69 immersion students and 22 native speakers in the context of the various written production tests were re-analysed from a lexical perspective, with verbs being selected as the focus for the study. The purpose of the study was threefold: (a) to compare different quantitative measures of immersion students' lexical proficiency in their second language (L2) writing; (b) to examine the relationship between written lexical proficiency and other aspects of their L2 communicative competence, and (c) to describe the students' lexical use in relation to that of native speakers.

(a) Measures of lexical proficiency. Five quantitative measures of lexical proficiency were developed and statistically compared. One of these was a 'lexical error rate', while the other four were variations on the theme of lexical richness, labelled respectively 'number of lexical types', 'lexical variety', 'lexical specificity', and 'lexical sophistication'. All the measures, except for 'number of lexical types' were controlled for length of text. For each student the data from the five written compositions were lumped together. Two of the relatively difficult measures were retained as the most appropriate for further use in a factor analysis. The first was 'lexical specificity', which consisted of the number of different verb types used by each student, not counting the 20 most frequent verbs in French or those that were used in the instructions to the compositions, divided by the number of verb items produced. The second measure was 'lexical sophistication', representing those relatively infrequent verbs not found in a basic word frequency list, also divided by the number of verb items produced.

(b) Lexical measures and L2 proficiency. Three mutually exclusive hypotheses arising from previous work were examined via factor analysis: (1) that lexical proficiency is equally involved in all three of the components of language proficiency examined in the large-scale proficiency study: namely, grammar, discourse, and sociolinguistics; (2) that lexical proficiency is part of the grammar component; or (3) that lexical proficiency is a separate component, distinct from the other three components of language proficiency.

Confirmatory factor analyses showed that an acceptable solution to fit any one of these three hypotheses could be found, and that there was no conclusive evidence favouring any one of the three hypotheses over the other two. One interesting finding was that in the three- and four-factor solutions corresponding to hypotheses (2) and (3) respectively, a grammar factor and a discourse factor emerged, which had not previously been found in the large-scale study where no lexical measures had been included.

(c) **Lexical use of immersion students and native speakers.** A comparison of the verb lexis used by the immersion students and the native speakers in their compositions revealed that the immersion students tended to make proportionately more use of high-coverage verbs, and less use of some morphologically or syntactically complex verbs such as pronominal and derived verbs. The inflectional complexity of some high coverage verbs did not appear to be a deterrent to their use although inflectional errors (considered grammatical rather than lexical errors) did occur. Semantic and syntactic incongruence with their English mother tongue (L1) emerged as an important factor in the immersion students' non-use of some French verb types and in the lexical errors they made. At the same time, the students demonstrated positive L1 transfer in the use of some cognate verbs in French.

Conclusions. It was suggested that the immersion students' stock of lexical items would benefit from more classroom activities designed to increase their use of L2 derivational resources and to emphasize the use of more specific vocabulary.

2:4 Communicative Skills of Young L2 Learners (Year 2 Report)

Purpose and data base. This exploratory study involved a detailed investigation of methods of scoring oral L2 performance and of the interrelationships among various aspects of L2 proficiency. The study was based on a subset of data previously collected in the context of another Modern Language Centre project. It consisted of oral tasks in English with 22 Japanese immigrant students in grades 2, 3, 5 and 6, together with academic tests of reading and vocabulary in the L2.

Findings. A comparison of global rating scales and detailed frequency scores as measures of specific aspects of oral L2 performance indicated that the two kinds of measurement were substantially correlated where there was sufficient variability in the data. An exploratory factor analysis of 26 variables, including measures of oral performance and academic test scores, yielded three orthogonal factors, interpreted as general English proficiency (including all the academic tests), vocabulary, and communicative style (consisting of interview variables). No separate factor was found for measures of fluency. Both the general English proficiency factor and the vocabulary factor were affected by length of residence in the L2 community, and general English proficiency was also affected by the students' age. Neither length of residence nor age was related to communicative style.

Conclusions. It was concluded that language proficiency results are strongly affected by the testing method (e.g. academic reading test, oral interview, story-telling task), and that an inherent difficulty in validating models of L2 proficiency is that measures faithfully reflecting a particular construct may not have adequate psychometric properties, while other psychometrically acceptable measures may fall short of representing the construct.

2:5 Metaphor Comprehension in Children's L1 and L2 (Final Report, Vol. I)

Purpose and design. This study compared the development of metaphor comprehension in Spanish-English bilingual children and monolingual English-speaking children, in order to test the hypothesis that metaphoric processing in bilinguals, as well as monolinguals, is constrained more by age and mental-attentional capacity than it is by language proficiency. Subjects were 20 Hispanic and 20 monolingual English-speaking children in each of three age groups: 7-8, 9-10, and 11-12 years, selected on the basis of a 'Figural Intersections Test' as being of normal mental capacity, which increased with age. An oral language proficiency test and a metaphor comprehension task in English were individually administered to each child. Hispanic children were also tested for oral proficiency in Spanish, and a subsample was tested for metaphor comprehension in Spanish. The language proficiency tests were similar to verbal IQ tests, while the metaphor comprehension task involved the oral interpretation of ambiguous metaphors, such as "my sister was a rock." The relative complexity of the children's metaphoric interpretations was coded with reference to the degree of semantic transformation involved in mapping an aspect of the vehicle (predicate) onto the topic (subject). The coding scheme had previously been shown to have good reliability and developmental validity for monolingual English-speaking children.

Findings. On the English proficiency test, Hispanic children scored significantly lower than the monolingual English-speaking children, and the Hispanic children resident in Canada for less than three years scored lower than those resident for more than three years. On the Spanish proficiency test, on the other hand, the more recent immigrants scored significantly higher than the long term residents. Performance on the metaphor comprehension task in English was, as predicted, found to be more strongly related to age and mental capacity scores than to oral language proficiency scores. While the bilingual Hispanic children did less well on the metaphor comprehension task than did the monolingual English-speaking sample as a whole, this was found to be related to the presence in the English-speaking sample of some students from a school in a higher socio-economic area. These children of middle-class background did better on the metaphor task than did the monolingual English-speaking children from the same schools as the bilingual children in working class areas. When the children of middle class background were removed from the sample, there was no main effect for language group on the metaphor scores, although the Hispanic children did less well on one of the two topics. Regression analyses indicated that the bilingual Hispanic children were similar to the subsample of English-speaking children from the middle-class neighbourhood in that English proficiency contributed little to the variance in their metaphor scores. Another finding was that conceptual structures developed in the first language appeared to facilitate metaphor comprehension in the second language, since for Hispanic children resident in Canada for less than three years, Spanish proficiency correlated more highly with metaphor scores in English than did English proficiency.

Conclusions. The findings of the study were in keeping with the hypothesis that, for bilingual as well as monolingual children, measured language proficiency was less predictive of metaphor performance than were age and non-verbal mental capacity scores. On a standardized test of English proficiency, the bilingual children scored significantly lower than their English-speaking schoolmates. On the metaphor task, however, the bilingual children performed almost as well as their English-speaking peers. This finding suggests that the metaphor task may be a more appropriate measure of conceptual skills in the second language than is a verbal IQ test.

3. CLASSROOM TREATMENT STUDIES

Several studies were undertaken to examine the relationship between instructional practices and the development of proficiency in a second language. During the first two years of the Project, a major focus was on the development and validation of a classroom observation instrument designed to capture the essential features of communication in the L2 classroom. This instrument was subsequently used in a process-product study which examined the impact on L2 proficiency of different instructional practices observed in core French classes. Two other studies grew out of the large-scale proficiency study described in 2:1 above. One of these involved the analysis of some specific aspects of language use and learning activities observed in French immersion classrooms, with a view to interpreting some of the earlier proficiency findings. The other study consisted of a classroom experiment in the French immersion setting, designed to enhance grammatical proficiency in the use of past tenses. These studies are summarized below.

3:1 Development and Validation of COLT Observation Instrument (Year 2 Report, Final Report, Vol. III)

The development of a new classroom observation scheme was motivated by the need to describe as precisely as possible some of the features of communication occurring in the second language classroom, and to distinguish between analytic and experiential orientations to language instruction. The COLT — Communicative Orientation of Language Teaching — scheme was derived from the communicative competence framework underlying the large-scale proficiency study and from a review of current issues in communicative language teaching.

Observation categories. The COLT observation scheme is divided into two parts. Part I, filled out by observers during the class, identifies different types of classroom activities and categorizes them in terms of: (a) participant organization (whole class activity, group work, individual work); (b) the content, or subject-matter, of the activity (e.g. classroom management, explicit focus on language form or function, other topics); (c) student modality (listening, speaking, reading, writing); and (d) materials in use (the type of material, length of text, intended users, and amount of control exerted on student language use). Part II of the COLT, which is later coded from a tape-recording of the class on a time-sampling basis, analyses communicative features of teacher-student interactions. Seven superordinate categories are identified: (1) use of target language (L1 or L2); (2) information gap (the level of predictability in an interaction); (3) sustained speech (length of utterances); (4) reaction to code or message; (5) incorporation of preceding utterances (how the participants react to each other's contributions); (6) discourse initiation (by teacher or student); and (7) relative restriction of linguistic form.

Validation. The observation scheme was piloted in 13 classes, mainly at the grade 7 level. There were four core French classes, two extended French and two French immersion classes, and five ESL classes in the sample. Each class was visited twice by two observers. Analysis of the Part I data entailed calculating the percentage of classroom time spent on the subcomponents of the various categories: participant organization, content, student modality, and materials. In the analysis of Part II, each verbal interaction feature was calculated as a proportion of its superordinate category. Results indicated that the COLT observation scheme was capable of capturing differences in the instructional orientation of the four types of classes. In core French

and ESL classes, for example, there was a heavier emphasis on form and more teacher control than in the extended French and immersion classes where there was a greater focus on meaning. Extended text was most often used in the immersion classes, and non-pedagogic materials were most often used in immersion and ESL classes. Sustained speech was least characteristic of the core French classes and most evident in French immersion and ESL classes. These comparative findings, intended as descriptive and not evaluative, generally met prior expectations concerning the various programs, except for some aspects of the ESL classes.

Conclusions. The ability of the COLT observation scheme to capture differences in instructional orientation was seen as an indication of its validity and as an important step toward identifying what makes one set of instructional techniques more effective than another.

3:2 The Core French Observation Study (Final Report, Vol. II)

In this process-product study, the COLT observation scheme was used to describe instructional practices in eight core French classes at the grade 11 level. Instructional differences were then analysed in relation to L2 proficiency outcomes in the different classes.

Subjects and procedures. The core French program was selected for study because the students' L2 proficiency could be assumed to derive largely from the classroom. The eight classes, from the metropolitan Toronto area, were preselected with the help of school board personnel to represent a range of L2 teaching practices. Early in their grade 11 year, the students were given a series of pre-tests of French proficiency, including some tasks from the large-scale proficiency study. The tests consisted of: (a) a multiple choice grammar test; (b) two written production tasks (a formal request letter and an informal note) which were scored for both discourse and sociolinguistic features; (c) a multiple choice listening comprehension test calling for the global comprehension of a series of recorded texts; and (d) an individual oral interview administered to a subsample of students from each class and scored for proficiency in grammar, discourse and sociolinguistics. During the school year, each class was visited four times for observation with the COLT scheme (in October, January, March and April). Observation periods lasted 40 or 70 minutes, depending on the duration of the class, and were tape-recorded. In May, the classes were post-tested with the same tests, and those students interviewed at the time of pre-testing were reinterviewed.

Analysis of COLT observations. Based on the Part I and Part II categories of the COLT observation scheme, it was possible to rank order the eight classes on a bi-polar composite scale from 'most experiential' to 'most analytic', based on the percentage of class time spent on practices defined as experiential in contrast to analytic. In the two most experiential classes, for example, there was proportionately significantly more topic control by students, more extended written text produced by the students, more sustained speech by students, more reaction (by both teacher and students) to message rather than code, more topic expansion by students, and more use of student-made materials than in the other classes. These two classes were labelled 'Type E' classes, in contrast to the remaining 'Type A' classes, where significantly more analytic features were in evidence, including a higher proportion of topic control by teachers, minimal written text by students, student utterances of minimal length, student reaction to code

rather than message, and restricted choice of linguistic items by students. The COLT analysis revealed at the same time that none of the classes was prototypically experiential or analytic, but instead intermediate along the bi-polar scale. The COLT findings were supported by teacher questionnaires providing information about classroom activities throughout the year.

The relationship of COLT findings to L2 proficiency. It was predicted that the Type A classes would be significantly higher on both written and oral grammatical accuracy measures than the Type E classes, but that the Type E classes would score higher on all other proficiency measures, including discourse and sociolinguistic measures, and scores on global listening comprehension. However, based on the post-test scores adjusted for differences in pre-test scores, no significant differences were found between the Type E and Type A classes, although a near-significant difference (p .06) emerged in favour of the Type A classes on the grammar multiple choice test. When the two Type E classes were compared to the two most analytic Type A classes (labelled Type A*), the Type A* classes did significantly better on the grammar multiple choice test (and specifically on agreement rules), but no other significant differences were found. A detailed correlational analysis relating the use of specific COLT features to L2 proficiency outcomes showed that there were both experiential and analytic features that were positively related to adjusted post-test scores. The profile of a successful core French classroom with respect to proficiency was identified as one in which a generally experiential approach was used with relatively more time devoted to features such as information gap, reaction to message, and topic incorporation. At the same time, positive correlations were found between a number of form-focussed, teacher-directed activities and proficiency outcomes. It was concluded that analytic and experiential approaches may be complementary rather than two ends of a continuum.

Qualitative analysis. In light of the fact that one of the two Type E classes made the most gain in overall proficiency and that the other Type E class made the least gain, the transcripts of these two classes were reviewed for qualitative differences that had not been captured by the COLT. The high-scoring class was found to engage frequently in communicatively rich interaction involving feedback and negotiation of meaning, while the low-scoring class received less feedback and spent more time on stereotyped routines. It was concluded on the basis of these findings that observation procedures based on COLT needed to be supplemented by more detailed forms of discourse analysis.

3:3 The Immersion Observation Study
(Final Report, Vol. III)

Classroom observations were carried out in nine grade 3 and ten grade 6 early total immersion classes in the Toronto and Ottawa regions, for the purpose of obtaining information on classroom processes and interpreting earlier immersion L2 proficiency findings. Each class was observed and tape-recorded for a full school day, including any instruction in English. All the tape-recordings were then transcribed. Analyses of some aspects of language use in the immersion classes are presented in the Project report. Further analyses are planned as time and finances permit.

Vocabulary instruction in immersion classes. L2 vocabulary-related classroom activities observed in the grade 6 immersion classes were analysed in the light of a classification scheme for describing such activities, and in relation to different kinds of linguistic knowledge involved in vocabulary learning. The analysis is qualitative rather

than quantitative. The classification scheme focusses on the following distinctions: (a) planned/unplanned instruction — the extent to which vocabulary instruction is seen as a purposeful activity; (b) systematic/haphazard instruction — the degree of systematicity with which specific features of vocabulary are taught; (c) written/oral activities — use of each medium for vocabulary instruction is seen to have a differential effect on lexical knowledge; (d) cross-linguistic/L2 based approaches to vocabulary teaching — a role for controlled use of the L1 is noted; (e) control of vocabulary selection — this may be by text authors, teacher or students; (f) the linguistic focus of teaching — whether the focus is on interpretation in context, conventional meaning, or other structural aspects of vocabulary. Linguistic aspects of vocabulary knowledge are categorized in terms of phonology, morphology, syntax, discourse semantics, graphology, and sociolinguistics. Analysis of the immersion classes in the light of these descriptors indicated that most planned vocabulary teaching occurred during reading activities organized around particular themes, during which students learned to pronounce words that they read aloud, to interpret passages, and in which the meanings of unfamiliar words were explained. Unplanned, spontaneous teaching of vocabulary was often student-initiated, as a specific word was needed to express an idea. The focus of both the planned and unplanned vocabulary teaching was mainly on interpretation and meaning. Given its association with reading activities, the teaching of new words emphasized written varieties of French and spelling. One example of how the students' prior knowledge could be drawn upon was provided by a teacher who drew attention to formal resemblances in the L2, enabling students to arrive at the meaning of an unfamiliar derived word. With some exceptions, the presentation of structural information about vocabulary was limited to the separate grammar lesson. Errors in gender, for example, tended to be only haphazardly corrected during other activities. There was no evidence that the teachers were focussing on sociolinguistic or discourse-related aspects of vocabulary. It was concluded that vocabulary teaching in the immersion classes occupied a rather narrow place in the overall teaching plan, and was mainly preoccupied with meaning and interpretation with insufficient planned attention to other aspects of vocabulary knowledge.

Vous/tu input. The underuse of vous as a politeness marker by early immersion students had been noted in the earlier proficiency study. The classroom observations provided an opportunity to relate these findings to use of vous and tu in the classroom context. Uses of tu and vous by the ten grade 6 immersion teachers and by the students in their public talk were counted and classified according to the functions they served: singular, plural, or generic; formal or informal. Teachers were found to use tu and vous about equally often, with tu generally being used to address individual children and vous to address the class as a whole. Occasionally, however, tu was used to the class and vous to individual students, leaving room for potential confusion. There was scarcely any use of vous by the teachers as a politeness marker, and its infrequency in this function in the classroom context was seen as a reason for its underuse as a politeness marker by early immersion students. Their underuse of vous in its plural function was at the same time attributed in part to the finding that, although used relatively frequently by teachers, very few opportunities appeared to arise for student production of vous plural in the classroom context. In conclusion, it was hypothesized that students would benefit from functionally-oriented instruction in the use of vous/tu, and opportunities to use them appropriately.

Student talk in teacher-fronted activities. It had previously been hypothesized that shortcomings in the grammatical competence of early immersion students may be due to

a lack of opportunity to produce 'comprehensible output', i.e. precisely conveyed messages demanding more rigorous syntactic processing than that involved in comprehension. In order to determine the opportunities that the immersion students had to talk in class, transcripts based on 90 minutes of French class time in each of the nine grade 3 and ten grade 6 classes were analysed, as well as the English portion of the day in the grade 6 classes. In general, the transcripts captured public talk rather than private, individual conversations. Each student turn was categorized according to length (minimal, phrase, clause, and sustained), and source (e.g., whether teacher- or student-initiated, preplanned or unplanned, linguistically restricted in some way or not, reading aloud from text, or reciting from memory). The findings indicated that in the 90-minute French portion of the day, student talk was less than two thirds as frequent as in the English portion of the day. Sources of student talk in French were very similar for the grade 3 and grade 6 students, the most frequent source being teacher-initiated student talk where the students' response was highly linguistically constrained, which appeared to encourage minimal responses from the students. Extended talk of a clause or more appeared to be encouraged when students initiated talk and when they had to find their own words. However, less than 15% of student turns in French were found to be sustained, i.e. more than a clause in length, when reading aloud was not included. It was concluded that greater opportunities for sustained talk in French by the immersion students are needed, and that this might be accomplished through group work, the provision of more opportunities for student-initiated talk, and through the asking of more open-ended questions by teachers.

Error treatment. An analysis of the grade 6 immersion teachers' correction of errors was based on the complete French transcripts of the ten classes observed. It focussed on the grammatical and pronunciation errors corrected by the teachers, the proportion of such errors corrected, and the systematicity of error correction. The highest proportion of error was observed in frequently used grammatical features such as gender, articles, and verbs. Only 19% of grammatical errors overall were corrected, but gender, article, and verb errors were more often corrected than other grammatical errors. About two-thirds of pronunciation errors were corrected. A lack of consistent and unambiguous teacher feedback was noted.

3:4 Functional Grammar in French Immersion
(Final Report, Vol. II)

This experimental study was designed to investigate the effect on immersion students' French proficiency of an approach to grammar teaching which involved the provision of focussed input in a problematic area of French grammar and provided students with increased opportunities for meaningful productive use of the target forms. Following a workshop with teachers, a set of classroom materials aimed at teaching the meaning distinctions between two major past tenses, the imparfait and the passé composé, were introduced for an eight-week period into grade 6 early immersion classes in six schools. These experimental classes were compared on pre-tests, immediate post-tests, and on delayed post-tests (three months later) with comparison grade 6 immersion classes in six other schools who were not exposed to the materials. The tests consisted of narrative compositions previously used in the large-scale proficiency study, as well as specially constructed cloze tests with rational deletions, and oral interviews administered to a sub-sample of students in each class. All the tests were designed to assess the students' ability to make appropriate use of past tenses and were scored

accordingly. There were two forms for each test, administered across testing sessions in a counterbalanced design.

The classroom materials. Adapted from an existing bank of activities focussing on the imparfait and the passé composé, the materials were divided into eight units, each to be used in a specific week. The teaching approach emphasized the integration of grammar teaching with worthwhile subject matter content and the personal experience of students. The oral and written activities, providing focussed input and opportunities for practice in using the two tenses, included the following: reading a simplified French-Canadian legend, discovering how the imparfait and passé composé served different functions in the legend, illustrating aspectually contrasting sentences, applying proverbs to the legend and to the students' own experiences, miming the progressive function of the imparfait, working in small groups to create new legends, and producing albums of childhood memories.

Findings. On the immediate post-tests, with adjustment made for pre-test scores, the experimental classes were significantly ahead of the comparison classes on two out of three measures: the cloze test and the oral interview. Three months later, however, at the time of delayed post-testing, there were no significant differences between the experimental and comparison groups on any of the tests. Both groups had improved their test performance over time. Evaluations of the materials by the experimental teachers at the end of eight weeks indicated general satisfaction with the materials, although some problems were noted with specific activities. Teachers indicated that they spent on average about 1 1/2 hours per week on the material. From some of their comments, it appeared that certain activities promoted more attention to subject matter content than to linguistic code, and informal observations in some classes indicated that past tense errors often went uncorrected during the 'Proverbes' activity. It was noted that one class with a teacher who was observed to provide frequent corrective feedback obtained the best results of all the classes on the composition test. Questionnaires administered to experimental and comparison group teachers at the time of the delayed post-testing indicated that the latter had also spent time working on the target verb tenses.

Conclusions. It was concluded that the teaching approach had succeeded in accelerating grammatical development in the experimental classes, but that to promote more long-term benefits some revision was needed in the materials, including more specific guidelines to teachers about the provision of corrective feedback. The fact that the control classes also appeared to have worked on past tenses was an additional factor that was surmised to have affected the long-term results.

4. SOCIAL CONTEXT AND AGE

The relationship between individual and social-environmental factors and the development of bilingual proficiency was examined in several minority and majority language learning contexts. In one large-scale study of Portuguese-Canadian students, the relationship between language use patterns, language attitudes, and bilingual proficiency was investigated by means of correlational and regression analyses, while in a small sample of beginning school-aged children of Portuguese home background, a detailed study of language interaction at home and at school was carried out with a view to relating interactional variables to later academic achievement. In another minority context, an ethnographic study focussed on students attending a French language

elementary school in Toronto. Finally, two studies examined the relationship between age and language learning: one among Japanese immigrant students of different ages and the other among Anglophone majority students learning French in three different school programs.

4:1 Language Use, Attitudes and Bilingual Proficiency of Portuguese Canadian Children (Final Report, Vol. III)

Purpose and design. In this study, the bilingual proficiency of grade 7 students from an important language minority group in Toronto was studied in relation to family background variables, the students' patterns of language use, and their language attitudes. Theoretical issues examined were: (a) the nature of language proficiency indicated by the pattern of relationships within languages; (b) the cross-lingual dimensions of language proficiency indicated by the pattern of relationships across languages; and (c) the extent to which proficiency in English and Portuguese could be predicted by language use and attitude variables.

The sample consisted of 191 students enrolled in Portuguese heritage language programs in seven inner-city Toronto schools. More than half these students were of Azorean background. The students all completed two questionnaires. One was a language use questionnaire concerning family background (e.g. birthplace, parents' language use, education, and occupations), language use patterns (use of Portuguese and English at home, in school, and in the community), and self-ratings of proficiency in English, Portuguese, and French. The other was a language attitude questionnaire which investigated dimensions such as integrative and instrumental orientations towards English and Portuguese, language use preferences in different contexts, the role of English and Portuguese in the students' ethnic identity, perceived attitudes of parents towards the students' education and language use, attitudes towards Portuguese dialects and language mixing, cultural assimilation, and attitudes towards French. Tests in English and Portuguese were also administered. In each school the students were divided randomly into three groups. One group did multiple choice grammar tests in English and Portuguese. A second group received a multiple choice discourse test in each language similar to the one administered in the large-scale proficiency study (see 2:1 above). Students in this group were also given individual oral tests in English and Portuguese, each of which contained tasks to be scored for grammar, discourse, and sociolinguistic proficiency. The sociolinguistic task in each language was adapted from the oral sociolinguistic test administered in the large-scale proficiency study. A third group of students in each school was given sociolinguistic written production tests in each language, again based on the test designed for the large-scale proficiency study.

The nature of language proficiency and its cross-lingual dimensions. A considerable degree of interrelationship was found among Portuguese self-ratings, multiple-choice discourse scores in Portuguese, and the various oral measures of Portuguese proficiency. A principal components analysis suggested a global Portuguese proficiency dimension, supplemented by academically related aspects of proficiency. Few relationships, on the other hand, were found among the measures of oral English proficiency, apparently because of a generally high level of performance giving rise to a lack of variability in scores. Across languages, self-ratings of proficiency in Portuguese, English, and French tended to be significantly related to each other. Further relatively strong cross-lingual relationships were observed for each set of written measures: i.e. between multiple choice grammar scores in English and Portuguese, between multiple

choice discourse scores in each language, and between written sociolinguistic scores in each language. These findings provided strong evidence for the interdependence of cognitive-academic skills across languages.

Predictors of bilingual proficiency development. Multiple regressions revealed that a considerable amount of the variance in the self-ratings of Portuguese proficiency could be related to attitudinal and language use variables such as students' acceptance of Portuguese, their knowledge and pride in Portuguese culture and achievements, their use of Portuguese media, exposure to Portuguese in the home, and their acceptance of and liking for French. Ceiling effects on the English self-ratings appeared to be at least partly responsible for the much weaker correlations found with attitude and use variables, although positive relationships were found with acceptance of English, use of English with siblings, and acceptance of French. The amount of exposure to Portuguese, both in Portuguese language classes and in the form of visits to Portugal, attendance at Portuguese mass, and Portuguese TV watching, appeared to be strongly related to measures of Portuguese proficiency, with weaker relationships noted between attitude variables and Portuguese proficiency. Minimal relationships were found between language use and attitude variables and the English proficiency measures, although there was evidence to suggest that positive attitudes towards Portuguese and students' use of Portuguese at home and in the community were in no way detrimental to their English proficiency.

Comparison with Azorean native speakers. A comparison of the Toronto students' test scores in Portuguese with those obtained by 69 grade 6 students in the Azores revealed that there were highly significant differences favouring the Azorean group on most measures of Portuguese proficiency. As in the large-scale proficiency study involving French immersion students (see 2:1 above), differences were most apparent on measures of grammar. The strong relationship found between Toronto students' attendance at Portuguese language classes and proficiency in Portuguese was seen as an indication that, in their minority context, more intensive exposure to Portuguese in an academic context could be advantageous for the bilingual development of the Toronto students.

4:2 Longitudinal Study of Young Portuguese Background Children: Bilingual Proficiency Development and Academic Achievement (Final Report, Vol. III)

Purpose and design. The major purpose of this ongoing study is to investigate the development of proficiency in both Portuguese and English in the transition from home to school. Twenty children from Portuguese backgrounds are being followed from the junior kindergarten year through grade 1 with respect to patterns of language interaction in the home, performance on a variety of language proficiency and literacy awareness measures, and (in grade 1) reading performance. Patterns of interaction in the home and knowledge of Portuguese and English will be used as predictors of English reading performance in grade 1. Thus, the study addresses theoretical issues such as the interdependence of L1 and L2 as well as practical issues related to the interaction between home and school variables in affecting the extent to which minority students are successful academically. The study also will provide a corpus of longitudinal data for analysis of students' developing proficiency in their two languages.

Methodology. The main sample consists of 20 Toronto students receiving the entire battery of tests. These are the Draw a Person Test, the Record of Oral Language (i.e.

sentence repetition) (English and Portuguese), Letter Identification (English and Portuguese), Concepts about Print (English and Portuguese) and, in Year 3 (Spring 1987), Test of Writing Vocabulary (English and Portuguese) and Gates McGinitie Reading Test (Comprehension subtest). (For test references, see complete study in Chapter 8.) In addition, children were taped in their homes for one and a half hours each year of the study.

Twenty-six grade 1 students (average age 7) in the Azores were also administered the Concepts about Print test, an oral interview, and Test of Writing Vocabulary (Clay 1979) in Portuguese for comparison purposes with the grade 1 Toronto data. In addition, six five-year-old students in the Azores were taped for one and a half hours in their homes. Data were also collected in Mainland Portugal from ten five-year-old children in a village situated a hundred kilometres northwest of Lisbon. A Portuguese version of the Record of Oral Language was constructed and administered to the children. Six of the ten were randomly chosen to be taped in the home.

Current status of the study. All the Year 1 home recordings have been transcribed and scoring schemes developed for grammar and pragmatics. A sample of students' transcripts have been scored but not the entire group. The Year 3 data will be collected in May and June of 1987. Subsequent to this data collection, a proposal will be developed to complete the transcription and data analysis relating home language use and proficiency in literacy-related aspects of English and Portuguese to English reading performance at the grade 1 level.

4:3 Ethnographic Study of a Toronto French Language School (Year 2 Report; see also Heller 1984)

In this ethnographic, sociolinguistic case study of a French-language elementary school in Toronto, patterns of language choice and language use were investigated in relation both to the micro-level interactional context and to the macro-level context of school and community. The study examined the role that the use of French and English played in the development of students' social identities.

Methodology. Micro-level data were collected in the school by means of participant observation over a six-month period, mainly in a grade 7/8 class, and through tape-recordings of eight students who each wore a tape-recorder for two entire school days. Four of the students were selected as ethnolinguistically representative of the school and the other four were randomly selected. Macro-level data were collected through a school-wide parent questionnaire and in interviews with school administrators, staff, members of the Parent-Teacher Association, and an ethno-linguistically representative subsample of parents.

Findings. Just over half the parents returned their questionnaires, which indicated considerable heterogeneity of family origins, linguistic backgrounds, and goals with respect to bilingualism and the maintenance of French. For example, over 40% of the families were of linguistically mixed marriages (usually with a francophone mother), 30% were francophone, 11% anglophone, and the remainder from a great variety of linguistic backgrounds. Very few parents and under half the children were Toronto-born. Family homes were widely dispersed over half of the city, making it hard for students to maintain friendships outside school. In-school observations revealed that there were three distinct groups of students: English-dominant, bilingual, and French-dominant.

The first two preferred to speak English among themselves, and the third -- a minority -- preferred French. Access to the different peer networks depended on appropriate language choice. Each group experienced its own tensions: French-dominant students reported pressure from peers to speak English outside class, while for English-dominant students, performance in French in class could be stressful. Bilingual students were observed to take part in occasional bilingual word-play and code switching, which was seen as their way of resolving the social tensions they experienced from their intermediate position and suggested that, for them, French and English were separate domains.

Conclusions. The heterogeneity of the school population and the varied linguistic experiences of the students were seen to militate against the formation of a monolithic French identity. Instead, observed patterns of language use indicated a close connection for the students between language choice and their evolving social identities.

4:4 Age on Arrival, Length of Residence, and Interdependence of Literacy Skills among Japanese Immigrant Students (Final Report, Vol. III)

Purpose and design. This study investigated the cross-lingual dimensions of language proficiency and the relationship between age and second language acquisition, with a focus on the development of reading and writing skills. We hypothesized that despite the dissimilarity of languages and writing systems, significant positive relationships would be found between Japanese minority children's L1 reading and writing skills and their acquisition of English reading and writing. An investigation of the relationships between Japanese and English proficiency appears to provide a stringent test of the interdependence hypothesis, which posits a common underlying proficiency for bilinguals, since the two languages have little in common at a surface structure level.

Subjects in the study consisted of 273 students between grades 2 and 8 attending the Japanese School of Toronto Shokokai Inc. Students were tested in May and June 1984 with measures of reading and writing in both Japanese and English. The reading comprehension subtest appropriate to students' grade level of the Gates-McGinitie Reading Test was given to all students who had been in Canada for at least six months as a measure of English reading skills. The Kyoken Standardized Diagnostic Test of Reading Comprehension published by the Research Institute for Applied Education in 1981 was given as the measure of Japanese reading skills. In addition, a letter-writing task in English and Japanese was administered to all children.

Scores on the English and Japanese reading tests were converted to T-scores to permit comparability across grades with the influence of age removed. In addition, English grade equivalent scores were used in some analyses as an approximate index of students' absolute level of English reading skills. A variety of indices of writing skills in Japanese and English were assessed.

Results. The results of correlational and regression analyses provide a consistent picture in relation to the acquisition of English reading and writing skills and their relationship to students' Japanese reading and writing proficiency. First, although the sample as a whole performs close to the mean (i.e. Japanese norms) in Japanese reading skills, there is a clear negative relationship between length of time in Canada and students' Japanese reading proficiency. The negative effect of length of residence on Japanese writing, however, appears minimal. Age of arrival in Canada appears to be a

more potent force in predicting maintenance of Japanese writing skills than length of residence. Similarly for Japanese reading, the older students are when they come to Canada, the better prospects they have for strong continued development of Japanese reading skills. This effect is not entirely due to the fact that students who arrive at older ages tend to have spent less time away from Japan, since the partial correlation between age of arrival and Japanese T-score remains significant even when length of residence is controlled.

It appears that students require about four years' length of residence, on the average, to attain grade norms in English reading skills. There appears to be some tendency for students who arrive at the age of 6-7 to make somewhat more rapid progress towards grade norms than those who arrive at older ages.

When length of residence is controlled, a significant relationship emerges between Japanese reading skills and English reading. Students' age of arrival in Canada (AOA) is also strongly related to English reading (controlling for length of residence), suggesting the influence of general cognitive maturity in mediating the cross-lingual relationship of cognitive/academic skills. General cognitive maturity, however, cannot account fully for the interdependence of reading skills across languages since significant relationships across languages were found for reading T-scores, in which the effects of age have been removed.

Writing performance was less closely related across languages than was the case for reading. This may be partly a function of the different types of measures used in each case (standardized reading tests v. non-standardized writing tasks). However, consistent significant relationships were obtained between Japanese writing and both English reading and writing measures. For some variables (e.g. Spelling) there was strong evidence of a specific cross-lingual relationship that was not mediated by more general cognitive/academic proficiencies.

Conclusions. In general, the data are consistent with previous studies in supporting the interdependence of cognitive/academic skills across languages. They also suggest that at least four years is required for students from highly educated backgrounds to attain grade norms on English academic tasks and that continued development of L1 academic skills to a high level (i.e. that of students in the home country) is a formidable task for students who arrive in the host country at an early age (particularly prior to formal schooling) but is considerably less problematic for students who arrive after several years of schooling in their home country.

4:5 Starting Age and Oral French L2 Proficiency in Three Groups of Classroom Learners (Final Report, Vol. III)

Purpose and design. The purpose of this study was to determine whether there are specific long-term advantages in oral L2 proficiency that can be associated with intensive L2 exposure at an early age in a total French immersion classroom setting. Three groups of grade 10 learners, with 11-12 subjects per group, were interviewed and given an oral sociolinguistic test in French: one group was from an early total immersion program which had begun in kindergarten, while the other two groups (from a late immersion and an extended French program respectively) had started their intensive exposure to French much later, in grade 7. A group of 12 native French speakers in grade 10 was also included in the study. The guided oral interviews were designed to

provide students with communicative contexts for the use of a range of verbs and verb forms. Transcripts of the interviews were analysed with respect to verb use and oral fluency in French. Scoring of verbs consisted of assessing the use of target verb forms in the context of specific questions, while the assessment of oral fluency was based on the nature and frequency of markers of disfluency and the linguistic contexts in which they occurred. The sociolinguistic oral test was based on the one used in the large-scale proficiency study (see 2:1 above).

Results. Group comparisons of the students' verb use indicated that the early immersion students were significantly more native-like on some variables (imparfait, conditional, use of pronoun complements in clitic position), but were no more native-like than the other learner groups on other variables such as use of number and person distinctions, time distinctions, and lexical variety, and in some instances tended to be less native-like than one or both of the other groups. The analyses of fluency revealed that in most types of disfluency, the three learner groups produced significantly more disfluencies than the native speakers but did not differ from one another. There was some evidence, however, that the early immersion students were producing fewer cut-offs and 'uh', 'um' etc. transition markers. The early immersion students were also less likely than the late immersion students to use transition markers in within-phrase locations, where such disfluencies were hypothesized to be more disruptive to discourse coherence than in between-clause or between-phrase locations. These findings indicated some advantages in oral fluency for the early immersion students who had started their intensive L2 program at a young age. Results on the sociolinguistic oral test, however, showed that the early immersion students did not manifest any general advantage over the other learner groups in sociolinguistic proficiency. While the early immersion groups displayed a slightly greater tendency to use attenuating conditional verb forms in formal social situations, they tended to be less sensitive to the appropriate use of the second person forms vous and tu than the late immersion and extended French students, whose intensive exposure to French in school had begun much later.

Conclusions. With respect to oral L2 proficiency, it appeared that there were some advantages to an early start in a French immersion program in the area of fluency and in the use of the verb system, but no advantage in the sociolinguistic domain. Some weaknesses in the verb system were also observed. As in other studies conducted in the early immersion context, a need for more emphasis on problematic areas in the target language system was indicated, along with greater opportunities for sustained oral and written expression.

5. CONCLUSIONS

The overall conclusions of the studies can be summarized with respect to the nature of bilingual proficiency and the influences on its development both in classroom and natural settings.

The nature of proficiency. At the inception of the study, the primary methodology envisaged for investigating the nature of language proficiency and its cross-lingual dimensions was confirmatory factor analysis. However, as a result of the findings of our Years 1-2 study of proficiency among French immersion students, in which little evidence emerged for the hypothesized trait structure, we became more explicitly conscious of the fact that the relationships between different components of language proficiency were a function of the specific language learning experiences to which

particular samples of individuals were exposed. This perspective implies a wider variety of analytic methods for investigating the nature of proficiency; specifically, we can discover a considerable amount about the nature of proficiency by observing its behaviour as a function of individual, social and educational conditions. Thus, we shall first consider the findings of our factor analytic studies and then examine findings of other studies that elucidate the nature of proficiency.

All studies that examined the relationships among different components of proficiency found significant correlations among written tests (including the core French observation study — see 3:2 above). These relationships were found across languages in the grade 7 Portuguese study (4:1), the Japanese study (4:4), and the metaphor comprehension study (2:5). Some evidence emerged for an oral factor (e.g. a communicative style dimension in the "Communicative skills of young L2 learners" study — 2:4) but the relationships among oral measures were considerably less strong than for the written measures. Similarly, some cross-lingual relationships among oral measures were found in the Portuguese grade 7 study but again the relationships were only marginally significant. These data are consistent both with the notion of a specific dimension of proficiency related to the ability to process language in context-reduced or decontextualized situations and with the hypothesis that this dimension is interdependent across languages.

There was considerably less evidence in the factor analyses for the hypothesized trait structure distinguishing grammatical, discourse, and sociolinguistic aspects of proficiency. However, discourse and grammar factors did emerge in the lexical proficiency study (2:3) and there was also some evidence for a separate vocabulary factor. The limitations of placing exclusive reliance on factor analysis for confirming hypothesized trait structures are illustrated in the fact that in this lexical study several mutually exclusive solutions produced an acceptable fit to the data. Also, in the original proficiency study (2:1), comparison of French immersion with native French speakers produced evidence that discourse skills were distinguishable from grammar and sociolinguistic skills, in that differences between L2 learners and native speakers were found only for the latter two aspects of proficiency.

Thus, consistent with the position advanced by Cziko (1983), the lack of strong support for the hypothesized trait structure in the factor analyses does not lead us to abandon the concept of traits. They are conceptually distinguishable and educationally important even if they are not statistically verifiable in relatively homogeneous school populations.

Classroom treatment. Our classroom treatment findings from different program settings lead to three main overall conclusions. First, there is evidence from both the core French and the immersion observation studies that the analytic focus and the experiential focus may be complementary rather than two ends of a continuum, and that they may provide essential support for one another in the L2 classroom. Second, the quality of instruction is clearly important in both analytic and experiential teaching. Analytic teaching will be successful in developing L2 proficiency only if it is appropriately matched to the learners' needs, while experiential teaching should involve communicatively rich interaction which offers plenty of opportunities for production as well as global comprehension on the part of the student. Third, learners may benefit if form and function are more closely linked instructionally. There is no doubt that students need to be given greater opportunities to use the target language.

Opportunities alone, however, are not sufficient. Students need to be motivated to use language accurately, appropriately, and coherently. In all these respects, the 'how' and 'when' of error correction will be a major issue for future investigation.

It seems reasonable to conclude that in all the programs under investigation — core French, heritage languages, and French immersion — much more work needs to be done in the area of curriculum design. Such work should include research to determine what combinations of analytic and experiential activities are most effective for different types of student. Another comparatively neglected area from the research point of view is teacher training and professional development. This area is likely to become more important at a time when more and more teachers are breaking away from their former dependence on prescribed pedagogic formulas and are increasingly making their own, more flexible, decisions about what can be done in the classroom.

Individual and social variables. With respect to the influence of individual and social variables on the development of proficiency, we can think of these effects in terms of the relative influence of attributes of the individual (e.g. cognition, personality) versus the target language input received by the individual. With respect to attributes, for example, it is clear from the Portuguese grade 7 and Japanese studies (4:1 and 4:4 above), as well as the immersion age study (4:5) that cognitive attributes of the learner play a significant role in at least certain aspects of target language acquisition. In the grade 7 Portuguese study and the Japanese study, children's cognitive/academic proficiency in their L1 was significantly related to the level of cognitive/academic proficiency attained in the L2. The relatively strong performance of late immersion students in comparison to those in early immersion is consistent with the notion that the learner's cognitive maturity (as indicated by age) is positively related to efficiency of L2 acquisition (at least up to the point where cognitive development reaches a plateau, possibly in the early to middle teens).

There is some evidence that cognitive attributes are more related to acquisition of certain aspects of proficiency than to others. For example, L1 cognitive/academic skills are more closely related in the Portuguese grade 7 study to performance on L2 written (context-reduced) tasks than is the case for oral tasks. Also, discourse proficiency appears to be somewhat less influenced by input/exposure variables than is the case for grammar, as illustrated by the native-speaker comparisons in the large-scale proficiency study (2:1) and Portuguese grade 7 study as well as in the regression analyses for Portuguese proficiency in the latter study (4:1).

In short, one way of thinking about the trait structure and its relationship to psychological variables is to distinguish between aspects of proficiency that are relatively more dependent on input from the environment for their full development than on attributes of the individual (e.g. oral grammar) and those that rely probably as much on individual attributes (e.g. cognitive skills, personality variables) as on input for their development (e.g. oral and written discourse, context-reduced proficiency generally). We would see sociolinguistic aspects of proficiency (particularly in the oral mode) as intermediate between grammar and discourse with respect to their relative dependence on input versus attributes. In the case of sociolinguistic proficiency, personality variables are likely to be at least as important as cognitive variables but input is clearly also crucial, as demonstrated by the immersion observation study (3:3), which showed minimal input to students regarding sociolinguistic variation. The relatively greater problems that early immersion students experience with grammar and sociolinguistic

proficiencies in comparison to discourse is consistent with this position, as is the more evident influence of exposure variables (e.g. visits to Portugal) on grammar than on discourse skills in the Portuguese grade 7 study.

In conclusion, the picture of bilingual proficiency that emerges from our studies is one of a dynamic evolving complex of traits that become differentiated from each other as a function both of variation in the input from the classroom or wider environment and the individual attributes of the learner.

Footnote

[1] In recognition that abstract, underlying language competence is not directly measurable, but inevitably coloured by the method of elicitation used, the term 'proficiency' is used in this report in a global sense to encompass both competence and performance aspects of grammar, discourse, and sociolinguistics that are measured by our tests.

References

Cziko, G. Some problems with empirically-based models of communicative competence. **Applied Linguistics**, 1983, 5, 23-38.

Heller, M. Language and ethnic identity in a Toronto French language school. **Canadian Ethnic Studies**, 1984, 16(2), 1-14.

Chapter 7

THE LANGUAGE USE PATTERNS, LANGUAGE ATTITUDES, AND BILINGUAL PROFICIENCY OF PORTUGUESE CANADIAN CHILDREN IN TORONTO

Jim Cummins, José Lopes and Mary Lou King

1. INTRODUCTION

The purpose of this study is to examine the bilingual proficiency of children from an important language minority group in Toronto and to determine the links with proficiency of family background, language use patterns, and language attitudes.[1]

The Portuguese community was selected for a number of reasons. They form a sizeable group in Canada and in metropolitan Toronto schools in particular they constitute the second largest ethnolinguistic group (after Italians). Moreover, a number of neighbourhood schools have sizeable proportions of Portuguese students and offer Portuguese classes as part of the heritage language program. Statistics regarding socioeconomic status show that social mobility is relatively low for this group as compared to other minority groups. Furthermore, the multiethnic survey conducted by Berry, Kalin, and Taylor (1977) showed that Canadians of diverse ethnic backgrounds view Portuguese Canadians as significantly less Canadian, less wealthy, and less similar to themselves in comparison to both English and French Canadians (see Table 1, p. 54). Finally, since French is quite similar to their native tongue, Portuguese speakers should exhibit favourable attitudes toward French and should learn French as a school subject quite easily.

Since the children were voluntary recruits from Portuguese language classes, they represent a sample of Portuguese background children which can be expected to be relatively favourably oriented toward Portuguese. Moreover, the fact that direct attitude measures, rather than indirect attitude measures, are employed in the study can also be expected to bias the findings toward responses more favourable to minority language varieties (e.g. see Bourhis 1986; Carranza 1982; Lambert 1967; Ryan, Giles, and Sebastian 1982).

Children in grade 7 are the focus of this study for several reasons. First, these youngsters are just beginning to be able to make their own choices regarding language use patterns and to evolve a sense of language/ethnic identity (Barker 1947; Day 1982; Lambert 1967; Ryan 1981). Up to this age, their language proficiency largely depends upon their family language background, the language use patterns around them, and the language attitudes of those around them. During the subsequent few years, they will be making important attitude-based decisions which will affect their further development of language proficiency and their future language use patterns. As bilingual adolescents, they will be facing choices in regard to their ethnocultural identity (Brody 1968; Derbyshire 1968; Lambert 1967).

1:1 Specific Goals of the Study

1. To investigate the relative proficiency attained in English and Portuguese for these children:
    - (a) Grammatical, discourse, and sociolinguistic aspects of proficiency for each language;
    - (b) Relation between self-ratings of proficiency and objective measures within each language;
    - (c) Relation between students' proficiency in Portuguese (L1) and English (L2);
2. To examine the relationships between family background (parental language background and socioeconomic status), age on arrival, and proficiency in the two languages;
3. To describe the relative frequency of English versus Portuguese language use patterns across different contextual domains with different interlocutors:
4. To describe attitudes toward English and Portuguese and perceived relative vitality of Portuguese Canadians in Toronto:
    - (a) The attitudes toward English and Portuguese (instrumental, integrative, preference, importance for identity);
    - (b) Perceived relative vitality of English and Portuguese;
5. To examine the relationship of language use patterns to proficiency (both objective and self-rated);
6. To examine the relationships of language attitudes and perceived vitality with proficiency;
7. To examine the students' relative proficiency in grammatical, discourse and sociolinguistic aspects of Portuguese in comparison to native speakers who have developed their Portuguese proficiency in a majority context.

1:2 Method

Subjects. The sample for the study consists of 191 grade 7 students (including 108 girls and 83 boys) enrolled in Portuguese heritage language classes in seven schools of the Metropolitan Separate School Board in the City of Toronto. Each of these schools offers a 30-minute daily program of Portuguese language and culture that is integrated into the school day. While the program is optional, it appears that the great majority of students of Portuguese background at the grade 7 level were enrolled in the program at the time of data collection.

The schools are located in an inner city area bounded roughly by Queen Street in the south and Dupont Street on the north, and by Spadina and Dufferin Streets on the east and west respectively. The area has a large concentration of residents of Portuguese background.

Participation in the study was voluntary, dependent on the students obtaining letters of permission from parents, and not all students in each heritage language class obtained permission to take part. Table 2, p. 55, shows the distribution of the sample across schools, and the proportion of students enrolled in the grade 7 Portuguese heritage language program at each school who participated in the study (on average 50 per cent). Also indicated in the table is the overall proportion of students attending each school in the relevant school year (1982-83) who were of Portuguese home background. Schools were selected that were relatively large in size (ranging from approximately 540-900 students), and that had at least two Portuguese heritage language classes at the grade 7 level. The different overall proportions of Portuguese background students in each school selected (from 60-90 per cent) reflects the range (55 - 95 per cent) in the Board's 13 schools which offer an integrated Portuguese heritage language program.

Of the total sample of 191 students, 25 (or about 13 per cent) were receiving special education. These students were not evenly distributed across schools in the sample. In two schools, the principals did not wish special education students to be tested, and one school had no special education program at the grade 7 level. The proportion of special education students in the sample from the remaining four schools is thus considerably higher (26 per cent).

Materials. Written tests of Portuguese and English grammar, discourse, and sociolinguistics, and oral tests in both languages were developed for the study. In addition, an existing oral test of Portuguese grammar and discourse was slightly adapted for use in the Toronto context. All tests were designed to fit the framework of traits and methods used in the French proficiency study reported in the Year 2 Report. In addition to the language proficiency tests, two questionnaires were designed: one concerning language use and the other, language attitudes. There follows a description of the various instruments, each of which was designed to be administered in a half-hour period. Copies of the questionnaires are appended to this report (see Appendix D, pp. 119-153).

1. Written grammar tests. Two multiple-choice written grammar tests were constructed for the grade 7 level: one in Portuguese and one in English. In the 50-item Portuguese grammar test, the students were required to select one grammatically correct sentence from each set of three sentences. The test assesses knowledge of a variety of grammatical features in Portuguese, including in particular, syntax (word order), use of prepositions, verb tenses, and verb agreement. The English grammar test consists of three parts containing items similar in nature to a number of standardized tests of English language skills designed for native speakers. (An earlier version of the English grammar test, similar in format to the Portuguese grammar test, proved when piloted to be much too easy for the target population.) In the first part of the test, which contains 15 items, the students have to indicate which sentence in each set of three contains a grammatical error, or if there is no error. In part 2, the students read ten independent sentences in each of which three segments have been underlined. Their task is to indicate which underlined portion of the sentence, if any, is incorrect. Both parts 1 and 2 of the test focus on a variety of grammatical

features in English, including verb agreement, past tense forms, and prepositions. Part 3 is concerned with syntax. It contains ten items involving groups of words (e.g. "were playing on the street"). The students are required to indicate whether the group of words is a complete sentence, or whether one of two choices (e.g. "the children" or "very happily") should be added to make a complete sentence.

2. Written discourse tests. Multiple-choice tests similar to the multiple-choice discourse test designed in French for the proficiency study in the immersion program setting (see Year 1 Report, pp. 11 and B27-B37) were constructed in Portuguese and in English to assess proficiency in recognizing coherent and cohesive text in the two languages. Some of the items from the original French test were adapted for use in either the English or Portuguese version; other items in each test were new. What was required of the students was that they identify the one sentence in a set of three which best fits the gap left in a short paragraph. Written instructions for each test were in both English and Portuguese to ensure that no student was impeded by lack of understanding of the unusual task to be performed.

3. Written sociolinguistic tests. Sensitivity to register differences in the two languages was assessed by means of written composition tests of a similar nature to the French sociolinguistic written production test used in the proficiency study (see Year 1 Report, pp. 15-16 and B53-B57). In the Portuguese test, students were required to write a formal letter of request to the Department of Tourism in Portugal requesting information about taking a holiday in Portugal. Fifteen minutes were provided in which to write the letter. Students then had a short (ten-minute) note to write, in which they played the role of a mother who is annoyed with her son/daughter instructing them to tidy up. In the English test, ten minutes were provided to write a formal letter of request to Environment Canada for information about pollution to be used in a class project. A note on the blackboard from the classroom teacher to a student who had neglected doing his/her homework was the next task to be performed (five minutes). A concluding section of the English test required students to rewrite an underlined portion in each of six paragraphs so that the register fits more appropriately with the passage as a whole.

4. Oral interviews. For each language, an individual oral interview was used to assess students' oral proficiency in grammar and discourse. For the Portuguese interview, the IDEA Oral Language Proficiency Test (Portuguese) (1981) was used. It consists of a controlled set of verbal and visual stimuli provided by the interviewer, to which the student responds. These responses were then assessed for correctness, appropriateness, and completeness. The interviewer checks responses on a test sheet as correct or incorrect. The test included items that focus on syntax, morphology (with a particular emphasis on verbs), lexis, and phonology. Three items also required discourse skills: (a) the production of two logical conclusions to a story provided by the interviewer; (b) a picture description task, and (c) a story-retelling task. A tape-recording of the interview was made so that some detailed scoring, particularly of discourse phenomena, could be done from the tapes. For the English interview, the original intention had been to use an English version of the IDEA test. However, like the written grammar test in English, this test proved when piloted to be too easy for the students. A new English interview was therefore constructed which involved more sophisticated conversational

interaction between the interviewer and student. Questions are nonetheless structured to permit assessment of control of verbs and prepositions, phonology, and vocabulary in English. Discourse tasks involve the telling of a story from a set of picture cartoons and the description of a telephone.

5. Sociolinguistic oral production tests. Two sociolinguistic oral production tests, one in Portuguese and one in English, were constructed, based on the oral test designed in French for the proficiency study in the immersion context (see Year 1 Report, pp. 12, 14, 15 and B14-B15). In each test, the student was shown a set of slides representing different social contexts. Each situation was accompanied by a taped description in the target language, and the student was required to produce the appropriate language in response to the slide and description. The students' responses were tape-recorded for subsequent scoring. Three sociolinguistic functions (request, complaint, and making an offer) were to be performed at two levels of formality (formal and informal). The purpose of the two tests was to determine the extent to which the students were able to adjust register appropriately in each of their two languages.

6. Language use questionnaire. The language use questionnaire consisted of 56 items concerning family background, language use, and language proficiency. Family background questions focussed on place of birth, location of schools, parental birthplaces, parental language background, parental education, and parental occupation. Language use questions concerned the extent to which Portuguese and English were used with family members as well as in school, in the community, and in leisure activities. For the traditional modalities (speaking, understanding, writing, and reading), students rated their proficiency for English, Portuguese, and any other language they know (usually French, which is given 20-40 minutes of instructional time per day). This questionnaire, available in English and Portuguese, was adapted from the language use questionnaire used previously by Mougeon with Franco-Ontarian students (Mougeon, Brent Palmer, Bélanger, and Cichocki 1982).

7. Language attitude questionnaire. The language attitude questionnaire was composed of 53 items, focussing upon five issues. Following Gardner and Lambert (1972), instrumental and integrative attitudes were distinguished and assessed for English and Portuguese. As well, direct preference questions (English versus Portuguese) were asked regarding which language is preferred in general and in specific contexts (e.g. expressing feelings or talking with friends). Four items addressed the role of the two languages in the respondents' ethnic identity. Adapted from the Parental Encouragement scale of Jakobovits (1970), the parental attitude towards language education scale assessed the perceived importance given by parents to the respondent's general education and to his/her learning of English and Portuguese in particular. Attitudes towards Portuguese dialects (i.e., mainland versus Azorean), language mixing, and code switching were addressed by eight items. Eleven items from the Subjective Vitality Questionnaire (Bourhis, Giles, and Rosenthal 1981) which seemed appropriate for Grade 7 students were adapted for the relevant target groups (Portuguese Canadians, English Canadians, and French Canadians). Four items assessed cultural assimilation in terms of an orientation to stay in Canada or return to Portugal and freedom of expression of opinions. Attitudes toward French were assessed in the parts of the questionnaire where the general lack of competence in that language did not make the item irrelevant (i.e.,

instrumental attitude, parental attitude toward language education, preferred language for school instruction). The questionnaire was available in both English and Portuguese.

Design. All students in the sample completed the language use and language attitude questionnaires. Students from each school were divided randomly into three groups (ensuring that one group, to be orally tested, always contained at least 12 students from each school). One group completed the written grammar tests in both languages, one group completed the written discourse tests, and the third group completed the written sociolinguistic tests. Students in the discourse group, which was the group with 12 students from each school, were also given all the oral tests. Table 3, p. 55, shows the number of students in the sample who completed each test.

Procedure. Each school was visited for a four- to five-day period by two researchers. The administration of the questionnaires and written tests took place in class on each of four days during the half-hour period devoted to the Portuguese heritage language program. On the first day, the language use questionnaire was administered, followed the next day by the language attitude questionnaire. Students were given the English version of the questionnaires, but the Portuguese version was available for any student who was more comfortable in that language. Help was provided to students, where necessary, with spelling of place-names, etc. On the third day, all English tests were administered simultaneously to subgroups of students in the class. And on the fourth day, the same subgroups received the Portuguese version of the written test that they had completed the day before (be it grammar, discourse, or sociolinguistics).

Oral interviews took place in a small room made available in the school throughout the school day, except during the heritage language class period. The twelve students writing the discourse multiple choice tests were interviewed on one day in Portuguese (by a male native speaker of Portuguese) and on a subsequent day in English (by a female native speaker of English). All interviews were tape-recorded. In one school, owing to the loss of recording equipment and seven students' interviews in English, a second female interviewer redid the seven interviews with these students two weeks later, and also carried out the English interviews with the remaining five students in that school.

2. SCORING PROCEDURES AND REFINEMENT OF THE DATA SET

This section outlines the scoring of the proficiency measures, the selection of indices, and the construction of composite variables. Oral Portuguese grammar, discourse and sociolinguistic measures are presented first, followed by written tests of Portuguese discourse, grammar and sociolinguistics. The six English measures are then discussed in the same order.

All proficiency variables analysed in the report are listed, with a brief descriptor of each, in Appendix A, pp. 85-87. Detailed scoring criteria are reported where appropriate in Appendix B, pp. 88-105, and reliability in Appendix C, pp. 106-118.

2:1 Oral Portuguese Grammar

Three tasks were scored to provide indices of students' grammatical competence in Portuguese. Two native speakers of Portuguese, referred to as scorer L and scorer R, participated in this and all subsequent scoring of Portuguese proficiency measures.

IDEA Oral Proficiency Test (Portuguese). Interviews were conducted by scorer L, who scored responses on the answer sheet at the time of the interviews. However, primary scoring was done from the tapes, after all testing was complete, by scorer L. (In the case of four items requiring a non-oral response - pointing or gesturing - the interviewer marked correct responses on the tape by saying 'mhmm'; when the tape and sheet disagreed, the record on the sheet was accepted.) Fifty-five short-answer items were scored.

Two items were discarded subsequent to the reliability check. Of the 53 retained, the proportion correctly answered was used as one oral grammar score (POGITOTL). Three subscores were computed, consisting of the proportion correct of (1) 17 vocabulary items (POGIVOCA); (2) 9 phonology items (POGIPHON); and (3) 27 remaining items dealing mainly with verbs, degree of adjectives, and prepositions (POGIGRAM).

Expressive language measures. A story-retelling task and a picture description task were embedded in the IDEA test.

The following were counted from transcripts of the students' story-retelling: (1) finite verbs; (2) five categories of grammatical error; (3) grammatical errors which could be considered as dialectal traits of Azorean or Brazilian Portuguese (subtracted from the total error count to create a second count of errors); (4) range of correctly used verb tenses (PVRBTNSS); (5) range of correctly used prepositions excluding 'a' (PPRPTYPS). Preliminary training and reliability checks were carried out, following which scorer L scored four classes, and scorer R, three classes.

For students with five or more finite verbs in the story retelling, two error rates were computed by dividing the total error count, and the error count minus dialect forms, by the number of finite verbs (POGERRAT, and POGEREXD, respectively).

The picture description was scored by scorer R for use of prepositions only. Three measures were obtained: (1) number of correctly used prepositions (preposition tokens; PPRPTOKP); (2) number of different prepositions correctly used (preposition 'types', a measure of range; PPRPTYPP); and (3) number of preposition errors (PPRPERRP).

Detailed scoring procedures for the expressive measures are reported in Appendix B.

2:2 Oral Portuguese Discourse

Scores for oral discourse competence in Portuguese were based (1) on the story-retelling task in the IDEA test that was also scored for grammar, and (2) on 'logical conclusions' produced to complete a short story (item No. 78-79 on the IDEA test.)

Story retelling. Three classes were scored by scorer L, four by scorer R. Detailed scores on a five-point scale (5 = high) were given for (a) logical sequence, (b) anaphora, and (c) time orientation. Following the detailed scoring, which was done from transcripts of the story-retelling, an impressionistic global discourse rating on a five-point scale was done directly from the tapes to allow any relevant non-linguistic oral patterns to influence the score. The discourse scoring strategies were similar to those originally designed for the study of grade 6 immersion students (see Year 2 Report, pp. 22-28). Further information on the detailed scores used in the current study is provided in Appendix B.

A composite score was computed as the mean of the scores for logical sequence, anaphora, and time orientation (PODDETL3).

Logical conclusions. Students were required to provide two conclusions to a story told by the interviewer (IDEA items 78-79), and were awarded 1 point for each conclusion that was judged to make sense, to a maximum of 2 points (PODCONCL).

2:3 Oral Portuguese Sociolinguistic Test

There were six sociolinguistic situations in this test: two called for the students to address a request to the person shown in the slide, two called for a complaint, and two for an offer. One offer was subsequently dropped because all students found the situation too difficult to understand. Detailed scoring of formal markers and a global impressionistic rating were carried out.

Detailed scoring. Detailed scores for the use of sociolinguistic markers were given on a three-point scale (3 = high, 2 = medium or mixed, 1 = low) for each of the following categories: initial forms of address, use of person in subsequent pronouns and verb agreement, and use of modality. A score on a two-point scale was also given for the presence (= 3) or absence (= 1) of extra politeness markers. The detailed scoring was done from transcripts. Further information concerning scoring criteria is provided in Appendix B, pp. 88-105.

Global rating. In addition to the detailed formality scores, an impressionistic rating of global appropriateness was given to each response on a five-point scale (5 = completely appropriate; 1 = completely inappropriate). The relationship between the detailed and global scoring was such that an informal variant which elicited no formal markers would be scored low on the detailed scores, but would tend to be given a high global rating if it was considered to be appropriate. The rating was done from the tape and the transcript simultaneously. Three classes were scored by scorer L, four by scorer R.

Composite scores were computed from both detailed scores and ratings of appropriateness. The four aspects (form of address, use of person, use of modality, and politeness markers) were the basis for one composite. For each situation the difference was computed between the score for each category of marker (score for formal variant - score for informal), and the mean difference for the situation was then calculated. The average of these difference scores (POSDIFF4), then, provided the first index of sociolinguistic competence. In addition, the average appropriateness rating for the formal variants (POSAPPHI) and for the informal variants (POSAPPLO) were computed, as well as the average of the two (POSAVAPP).

2:4 Written Portuguese Discourse (Multiple Choice)

For the multiple choice test of discourse in Portuguese, students were presented with 30 three-choice right-wrong items. On each, they were required to select the sentence which best fit in the blank space in a short paragraph. Four items were deleted from the analysis because of typographical errors. The proportion correct of the remaining 26 items (PWDMULTC) was used as an index of competence in written Portuguese discourse.

2:5 Written Portuguese Grammar (Multiple Choice)

The multiple choice test of Portuguese grammar consisted of 50 items, one of which was deleted because of a typographical error. All items were three-choice, and scored as right or wrong. The proportion correct out of 49 (PGRAMMC) provided a measure of competence in written Portuguese grammar.

2:6 Written Portuguese Sociolinguistic Test

The request letter to the Portuguese Department of Tourism and the mother's note were each given detailed scores for the use of formal sociolinguistic markers. The assumption was that the liberal use of formal markers in the request letter would be sociolinguistically more appropriate than their use in the note. An impressionistic global rating for 'sociolinguistic appropriateness' was also awarded for each task.

Detailed scores. A three-point scale was used to assign formality scores (3 = high, 2 = medium, 1 = low) in each of the following categories: the opening of the letter or note; the use of person in subsequent pronouns and verb agreement; the use of modality; and the closing of the letter or note. An additional score on a two-point scale was given for the use of extra politeness markers; in this instance, a score of 3 indicated the presence, and a score of 1 the absence, of formal markers. See Appendix B for further information about each of the above categories.

Global rating. A five-point scale (5 = completely appropriate, 1 = completely inappropriate) was used in rating the 'sociolinguistic appropriateness' of the language used in each task. In this global impressionistic form of assessment, a high rating of 5 given to a note would indicate not that it contained numerous markers of high formality, but that the language used was judged to be completely appropriate in the given context.

From the above, three indices of sociolinguistic competence in written Portuguese were derived. A difference score (PDIFFORM) was computed from the detailed scores by subtracting the sum of the scores on the note from the sum of the scores on the letter. Two additional indices were based on the global ratings: the rating for the letter alone (PLETGLOB) and the average of the ratings for the letter and the note (PGLOBAV).

2:7 Oral English Grammar

Scores for oral grammatical competence in English were based on responses to specific questions that students were asked in individual conversational interviews. The interviews were tape-recorded and transcribed later. Unless otherwise noted, scoring was done from the transcripts. Scores were given for the use of prepositions and verbs, for syntax, and for pronunciation. For more detailed information regarding scoring criteria, see Appendix B.

Use of prepositions. Each student's response to question 14 (...Tell me the story the pictures tell...) was scored by a native English-speaker for preposition use. Counts were made of (1) the number of preposition errors made, (2) the number of obligatory contexts for prepositions, (3) the number of different prepositions correctly used (types), and (4) the number of prepositions actually used (tokens). From these, two measures were derived: a preposition error rate (number of preposition errors/number of obligatory contexts - .EPRPERAT), and sophistication of preposition use (number of different prepositions used correctly/total number correctly used - EPRPTYTO). Any student for whom the denominator of either ratio was less than five was treated as having missing data for that variable.

Use of verbs. Responses to three questions were scored for errors in use of verbs. The questions were:

7. What are you going to do today after school?

8. Can you tell me three things you did before you came to school today?

9. What would you do if you won a million dollars in a lottery?

A count was made of the number of verb errors and of the number of obligatory contexts for verbs. The ratio of errors to obligatory contexts was computed (EVRBERAT) for students with at least five obligatory contexts.

Lexical sophistication of verb use was scored on the story telling task. A type/token ratio (EVRBTYTO) was computed from the number of different verbs correctly used (types) and the number of verbs produced (tokens) for students providing at least five tokens as a data base.

For both measures of verb use, two native speakers of English either did joint counts (lexical sophistication) or divided the scoring equally between them and then checked each other's scoring (verb errors). No reliability is, therefore, reported.

Syntax. Sophistication of syntax was assessed in terms of average number of words per 'Communication-units' or 'C-units' (ECUNITSZ) used in the story telling. A C-unit is defined as a clause plus its modifiers. Again, the two scorers worked together on assessing this measure; no reliability is reported.

A second measure of syntax was also based on the story-telling task. Working from the tape (after scoring the story for Discourse), the raters made an impressionistic global judgement of oral syntax on a scale from 1 (intelligible only in set expressions) to 5 (native-like). The final global score for oral syntax (EORLSYNT) is the mean of the two raters' scores.

Pronunciation. After judging the story-telling task for syntax, the raters judged the student's pronunciation impressionistically on a scale from 1 (intelligible only in set expressions) to 5 (native-like). The final score for pronunciation (EPRONUNC) is the mean of the two scorers' judgements.

2:8 Oral English Discourse

Oral English discourse competence was scored on the tapes of the story-telling task. Detailed criteria for scoring are to be found in Appendix B.

Global ratings. Scoring was done by each of two raters, scoring directly from the tape. The rater first listened to the tape and judged oral discourse globally on a scale from 1 (completely disjointed) to 5 (very well connected). After each five to ten students, they compared their ratings and discussed any major discrepancies. The mean of their two ratings was computed (EODGLOBL) for use in analysis.

Detailed scoring. The tape was not replayed unless necessary; detailed scoring was carried out for each student immediately following the global rating. A rating from 1 (low) to 5 (high) was assigned on each of the following four details: (1) setting the scene and identification of characters (EODSCENE); (2) logical sequence of events (EODLOGIC); (3) anaphora (EODANAPH); and (4) logical sentence connectors (EODCNECT). As for the global scoring, the scorers compared their assessments at regular intervals. The final score on each detail is the mean of their two judgements. In addition, an overall score for each student was computed as the average of the four detail scores (EODDETL4).

2:9 Oral English Sociolinguistic Test

The form of the test for English oral sociolinguistic competence exactly parallels the Portuguese version; only the content of the six situations differs. Because of differences between the two languages, the precise details scored differ. The scoring was done by one English native speaker; the procedure involved scoring all students for one item at a time.

Each item was scored for the use (= 1) or non-use (= 0) of each of eight categories of markers or formality; the score for each item was the number of categories used. The categories were:

1. interrupter (e.g. "pardon me")
2. form of address (e.g., "sir", "missus")
3. politeness markers (e.g., "please"
4. modal auxiliaries in the main speech act
5. grammatical mood in the main speech act
6. conditional clause
7. attenuation
8. explanation

The categories are fully explicated in Appendix B, pp. 88-105.

After detailed scoring was complete, the transcripts were re-read, and a global impressionistic judgement from 1 (inappropriate) to 5 (appropriate) was made of the response to each item.

To maintain comparability with the Portuguese test, one item (requiring the student to offer to go to buy a pen for a teacher/friend) was excluded from the analyses.

Composite scores were computed for both detailed and global measures:

1. mean detail score across the five low formality items (EOSDETLO)
2. mean detail score across the five high formality items (EOSDETHI)
3. mean difference score (EOSDETHI - EOSDETLO) across the five items (EOSDIFF)
4. mean appropriateness for the low formality items (EOSAPPLO)
5. mean appropriateness for the high formality items (EOSAPPHI)
6. mean overall appropriateness (EOSAPPHI + EOSAPPLO)

2:10 Written English Discourse (Multiple Choice)

As a multiple choice test of discourse in English, students were given 32 three-choice, right-wrong items. On each, the student had to select the sentence which best fit in the blank space in a short text. No item was deleted before analysis. The score was the proportion correct of the 32 items (EWDMULTC).

2:11 Written English Grammar (Multiple Choice)

The multiple choice test of English grammar consisted of 35 items, one of which was dropped because adult English native speakers failed to agree on which option was the correct answer. All items were scored as right or wrong. Items 1 to 25 were four-choice; 26 to 35 were three-choice. The score was the proportion correct of the 34 items retained (EGRAMMC).

2:12 Written English Sociolinguistic Test

There were three sections to the written sociolinguistic test in English. Students had to compose a letter to Environment Canada requesting information for a project, and a note on the blackboard as from a teacher to a negligent student. They also had to rewrite short segments in each of six paragraphs so as to make the register fit the rest of the paragraph more appropriately.

Letter and note. A global rating was made of the body (excluding opening and closing) of the letter and of the note. The scale was from 1 (inappropriate) to 5 (appropriate). Both the letter and the note were also scored for appropriateness of (1) the opening/closing (on a three-point scale) and (2) mitigation of the directive (on a five-point scale). Since two raters scored all compositions, the mean of their two scores was used on each measure.

Following the reliability study and exploratory analyses, the rating of appropriateness of mitigation in the letter (ELETMIT) and the average of the global appropriateness ratings for the letter and note (EGLOBAV) were selected for analyses.

Rewriting task. On the rewriting task, each of the six items was rated impressionistically by three scorers for the extent to which it improved on the original phrasing. Two items (no. 1 and 5) were originally scored on a three-point scale; values were later rescaled to 1, 3 and 5. The other items were scored on a five-point scale. The score for each item was the mean of the three scorers' ratings; if one of the three coded the item as missing or unscorable, the mean of the other two judgements was taken. The final score (REWRITE) used in the analyses was the mean of the six item scores.

For further details on scoring criteria for mitigation and appropriateness of the letter and note, and appropriateness of the short answer items, see Appendix B.

3. RESULTS

The results of the study will be described with respect to the three major theoretical issues to which the study is addressed: (a) the nature of language proficiency as indicated by the pattern of relationships within languages; (b) the cross-lingual dimensions of language proficiency as indicated by the pattern of relationships across languages; and (c) the predictors of bilingual proficiency development as indicated by the relationships between Portuguese and English proficiency, on the one hand, and use and attitude variables on the other. Initially, however, in order to provide a context for consideration of these theoretical issues, the general trends that emerged from the Use and Attitude questionnaires will be described.

3:1 Students' Language Use and Attitude Patterns

Descriptive statistics for the Use and Attitude Questionnaire variables are presented in Appendix D, pp. 119-153, together with the questionnaire items themselves. Only the most salient patterns are described here.

Language Use Questionnaire. Part A of the Language Use Questionnaire provides information about the place of birth as well as on the urban or rural background of each respondent. As could be expected among a Portuguese speaking population in Toronto (even though there are no official statistics to support it), the largest group of respondents (40%) was of Azorean origin. In addition, 35% were born in Canada, 15% in Mainland Portugal, 4% in the Madeira Islands and 7% elsewhere (mainly the former Portuguese Colonies). Of the total population, 66% were born in an urban setting and 34% in a rural setting.[2]

Part B was directed at students not born in Toronto. They were asked how old they were at the time of arrival in Toronto and where they had lived previously. Again, in Question 5 (Q.5) the majority of respondents (56%) indicated that they had lived in the Azores before coming to Toronto. However, as judged by the scorers, 65% of the respondents indicated they had lived in a rural setting and 35% in an urban environment.

Of the 11% that had lived in places other than the ones indicated in Q.5, only 10% of these had lived in other parts of Canada compared to 90% that had lived in other parts of Portugal.

In summary, the largest proportion of respondents were from the Azores where they had lived mainly in rural settings.

In Part C, students were asked with whom they were living in order to determine if they were part of an extended or a nuclear family. The latter, as far as the Language Use Questionnaire is concerned, is considered a family that includes only parents and siblings. Students were also asked about the birthplace, formal education, occupation and native language of both parents. Questions related to the parents' formal education and occupation were included in order to determine their socio-economic status. With respect to type of family in which the respondents were living (see Q.8), it can be observed that the majority were living with both parents as well as with brothers and sisters, while 26% lived in an extended family that included grandparents or others. Others, in this case, could refer both to other family members or to possible tenants living in the same house. The majority of both fathers and mothers were born in the Azores (60% for each); no parents were born in Canada. The father's birthplace was coded as rural in 72% of cases and urban in 28%; the same distribution held for mother's birthplace.

With respect to parents' formal education, the majority had attended only primary school (73% for the fathers, 76% for the mothers); 14% of the fathers and 8% of the mothers never attended school and only 2% of both parents attended university or college. Answers dealing with the parents' occupations were coded, whenever possible, using the **Occupational Classification Manual, Census of Canada** (1971). It is clear that the majority of parents could be considered as belonging to the working class. The native language of the parents was Portuguese in 99% of cases.

In Part D students were asked about the schools they had attended in order to determine what previous language exposure they had received in a school setting. They were also asked to rate themselves on a five-point scale on the amount of spoken Portuguese used before going to school, and on their abilities in spoken and written Portuguese, English and French at the time the Language Use Questionnaire was administered (i.e., when they were in Grade 7). Another set of questions dealt with the specific patterns of language use in and outside the home.

With respect to previous schooling, the students were exposed primarily to English (94%), although 4% reported primary exposure to Portuguese in the school they had attended for the longest period of time. Almost all the schools were in urban settings which, together with the amount of exposure to English, suggests that relatively few of the respondents attended school before coming to Canada. Concerning the language they spoke before going to school one can see that the majority (78%) spoke Portuguese either 'often' or 'always'. However, when asked to rate themselves concerning current understanding, speaking, reading and writing, the respondents rated English much higher than Portuguese (see Qs.24 to 31). This means that by Grade 7 the students considered themselves to be considerably more fluent in English. As far as their fluency in French was concerned, the students rated themselves slightly below Portuguese (see Q.32).

Concerning the language they used to speak to their family, relatives and neighbours, the students indicated that Portuguese was mainly used with the older members of the family, i.e., with grandparents. Portuguese was also mostly used to speak with the parents; however, when speaking to brothers and sisters English was mostly used. The same applies to the language used by family and relatives (see Q.34) when they addressed the respondents. The students reported that the parents always used Portuguese (100%) to speak with each other, which means, together with the information in Qs. 33 and 34, that the respondents were receiving high exposure to Portuguese in the home; yet, English was practically always used when the respondents spoke among friends (98% English, 2% Portuguese).

In Part E students were asked about other situations in which they used language or were exposed to it, such as television, radio programs, books they read, and letters they wrote to friends and relatives. They were also asked if they attended Portuguese classes in the Portuguese community clubs and about their language use in other activities such as folk dancing classes. The respondents watched a considerable amount of television. In fact, they watched Portuguese programs on television for an average of 2 hours a week and English programs for an average of 23 hours a week. As far as books the respondents read, 64% read hardly any books in Portuguese, yet the majority of them read books in English (35% three or more books a week, 27% one/two books a week and 24% two/three books a month). With respect to radio programs, 31% never listened to them in Portuguese while 36% listened to Portuguese radio only once in a while. In English, however, 34% reported listening to radio programs an average of four hours or more every day. The respondents wrote somewhat more letters to friends and relatives in English as compared to Portuguese; 45% never wrote letters in Portuguese while 30% never wrote them in English. Twenty-one percent of the respondents attended Portuguese classes for several years at the clubs in the Portuguese community for an average of ten hours a week; 16% were involved in a Portuguese folk dance group where the dancers used Portuguese and English about equally while instructors used Portuguese more than English (see Q.48). Concerning other community activities it may be noted that although the majority attended mass in Portuguese, in other activities, such as special parties and/or festivities, the majority used Portuguese and English about equally (see Qs.49 to 55). English was always or mostly used by the vast majority of respondents at school, during sports activities and among peers, a pattern which is consistent with the results obtained in Qs.33 and 34. With respect to the number of trips to Portugal within the last five years, 50% had not been there, 25% only once, 11% twice, and 13% had been there three or more times.

From the Language Use Questionnaire it can be inferred that the respondents, the majority of whom were of rural Azorean background, were exposed primarily to Portuguese in the home, especially at a younger age (i.e., before going to school). After the beginning of school, a rapid transition to English appears to have taken place, with the result that with friends, as well as with brothers and sisters, English was the language used almost all the time by these grade 7 respondents. For most students, use of Portuguese was restricted to Portuguese classes (in both the Heritage Language Program and at the various Portuguese classes in the clubs), to mass and some other activities of the Portuguese community, to address older members of the family (i.e., grandparents) and, somewhat less frequently, to communicate with their parents.

Language Attitude Questionnaire. The principal aim of the Language Attitude Questionnaire was to inquire about students' language preferences and attitudes towards the Portuguese, English and French languages. The questionnaire also contained a set of questions on ethnic identity, attitudes towards dialect and language mixing, and a set of questions on the parents' attitude(s) towards language education.

In Part A students were asked to rate on a five-point scale the importance of Portuguese, English and French as it pertained to their academic/professional goals, and the relevance of Portuguese and English to communicate with family, friends and people in general. In the questions concerning academic/professional goals (i.e., Q.1, "to do well in school"; Q.6, "to get into a good high school"; Q.4, "to get into a good university"; Q.8 "to get a good job in the future" and Q.10, "to be a well educated person"), the majority of respondents rated all three languages as being important, although English was rated as considerably more important than either Portuguese or French. In questions related to the importance of Portuguese and English "to have an important position in the Portuguese Community" (Q.2) and "to communicate with my family" (Q.3), Portuguese is rated much higher than English (i.e., for Q.2, 63% of respondents answered that Portuguese was extremely important as opposed to 36% who answered that English was extremely important; similar results can be seen on Q.3: Portuguese 74%, English 21%). Yet in questions where they are asked about the importance of both languages "to make good friends" (Q.6) and "to be more comfortable with people around me" (Q.9), English was rated much higher than Portuguese, especially on Q.6 where 79% of respondents answered that English was extremely important as opposed to 13% who accorded the same importance to Portuguese. The language attitudes revealed in this part of the questionnaire seem to relate to certain domains of language use as described in the summary of the Language Use Questionnaire data. For example, students used Portuguese in the home, especially to address older members of the family, yet they preferred to use English when talking among friends.

In Part B the students were asked with respect to Portuguese, English and French, which language they felt comfortable speaking (Q.11), which language they liked (Q.12), in which language they would like to be taught (Q.13) and in which language they liked to watch television (Q.14). On this set of questions, they rated English much more highly than Portuguese, while French was rated below Portuguese. On Qs.15 to 19, students were asked on a five-point scale that went from "Portuguese all the time" to "English all the time" which language they preferred to use to talk about feelings (Q.15), to talk with friends (Q.16), family (Q.17) as well as which language they preferred their families to use (Q.18) and the language they preferred to use in talking about things that happened in school (Q.19). Again, respondents showed a preference for talking to their families in Portuguese while they preferred to use English with friends or when talking about their own feelings.

In Part C the students were asked a set of questions concerning their linguistic identity (Qs.20 to 23). Although they preferred to use Portuguese primarily to talk with their families, the majority of respondents considered themselves as speakers of English and Portuguese; they showed a desire to become, in the future, speakers of Portuguese, English and French, to marry Portuguese/English speakers and to have their children be fluent in both Portuguese and English.

Part D consisted of a set of statements concerning the vitality of the Portuguese community. Students were to rate these statements on a five-point scale, from "strongly agree" to "strongly disagree". The majority of respondents considered that Portuguese should be kept alive in Toronto, although they felt that in Toronto and in the world people recognized the value of English more than Portuguese and French (Qs.24 to 27). They indicated that among English-, Portuguese- and French-Canadians, English-Canadians were the most proud of their history, language and culture (Q.30), that in Toronto they constituted the most recognized group (Q.29) and that English-Canadians thought the study of French was more important than the study of Portuguese (Q.28). They were aware that the Portuguese community in Toronto had hardly any contacts with French-Canadians and considered that the community established more contacts among its own members than with English-Canadians (Q.31).

In Part E, students were asked about the quality of Portuguese and English spoken in their own homes as well as their attitude towards dialects (i.e., Azorean vs. standard Portuguese) and language mixing. The majority of respondents either agreed or strongly agreed that Portuguese spoken in their homes was very good. However, they had mixed feelings concerning the quality of English (see Qs.32 and 33). Regarding their attitudes towards dialects there were no clear-cut results since a large number of respondents had mixed feelings. Yet a large number agreed that there might be communication problems across dialects. Regarding the mixing of Portuguese and English, students were somewhat more tolerant of using some English words when speaking Portuguese than vice-versa. With respect to the dialect used to teach Portuguese at school, the majority of respondents did not show any specific preference (see Q.39). The breakdown was as follows: Azorean-Portuguese 15%, Mainland Portuguese 23%, both Azorean and Mainland Portuguese 21%, and "it does not matter" 41%.

Part F consisted of a set of questions regarding parents' attitudes towards higher education and the interest shown in their children's studies. The large majority strongly agreed that their parents showed an interest in their studies and would like them to have an academic education and go to university (see Qs.40 to 42), yet they felt their parents had less regard for French than for English and Portuguese (see Qs. 43 to 45).

In Part G, students were asked about how comfortable they felt about expressing their own opinions with family, friends and adults in general as well as about their plans to stay in Canada and how they felt about Portuguese and Canadian cultures. As far as expressing their opinions was concerned, they felt more comfortable expressing these with parents, siblings and friends than with teachers or adults outside the family (see Q. 46); 87% answered that they planned to stay in Canada and 73% felt their families planned to do the same (see Qs.47 and 48). Regarding their knowledge of Portuguese and Canadian history and achievements and how proud they felt about these (Qs.50 to 53), the respondents showed similar feelings about these cultures.

In summary, the picture that emerges from students' responses to the Language Attitude Questionnaire is one of a generally realistic appraisal of the status of Portuguese and English for both their community and the wider population. Although English is viewed as considerably more important for activities outside the Portuguese community, the importance of Portuguese within this community is recognized and valued by most of the students themselves. Most students appear to be comfortable with their dual identity as Portuguese-Canadians.

Conclusion. The descriptive data from the Language Use and Language Attitude Questionnaires reveal, on the one hand, the strength of students' involvement and identification with mainstream Canadian culture, insofar as this is represented by their peers and the wider social institutions in which they participate (e.g. schools). English is used almost exclusively among siblings and friends, although many of these are also from Portuguese backgrounds. Despite the fact that Portuguese is still the primary language used by students at home, they rate their proficiency in Portuguese considerably less highly than their proficiency in English.

These patterns of use reflect students' perceptions of the relative status of Portuguese and English in Toronto. Although students are generally positively oriented towards Portuguese, they recognize the greater status of English and the importance of proficiency in English for future educational and job-related success. There is little evidence, however, of rejection of their Portuguese identity in favour of an English-Canadian identity. Most students appear to have achieved a relatively comfortable identification with both cultures. It should be noted, however, that these data reflect only one point in a developmental process and that at earlier ages students may have had more negative feelings towards Portuguese, a possibility that may have influenced their relative skills in each language.

3:2 Relationships Among Proficiency Variables Within Languages

Portuguese proficiency. A large number of potential indices of grammatical, discourse, and sociolinguistic skills was derived from the different tasks administered to the students. The inter-relationships among these indices were examined both to assess the extent to which different indices reflected a unitary or complex trait and to select the most appropriate indices for more detailed analyses. Descriptive statistics, including means, standard deviations, and number of cases for these variables for all students in the orally tested sample are presented in Table 4, p. 56.

The correlations among the oral Portuguese grammar variables are shown in Table 5, p. 57. Relatively low correlations were found among the indices derived from students' story retelling and picture descriptions. However, the total score on the IDEA test (POGITOTL) correlated significantly with the majority of indices derived from the expressive language measures. The total IDEA score was also broken down into grammar (POGIGRAM), vocabulary (POGIVOCA), and phonology (POGIPHON) indices, and grammar and vocabulary were found to show similar patterns of relationships with other grammar variables as POGITOTL. The phonology index, however, tended not to correlate highly with most grammar variables.[3]

The correlations among oral and written Portuguese discourse indices are shown in Table 6, p. 58. Moderately significant relationships are observed both between written and oral discourse indices and within the oral indices. The global rating index (PODGLOBL) was selected as the best overall measure of oral Portuguese discourse for subsequent analyses.

The correlations among oral Portuguese sociolinguistic indices are presented in Table 7, p. 58. As with Portuguese discourse measures, correlations among different indices are relatively high. The difference between 'high' (formal) and 'low' (informal) situations in students' use of four categories of formality markers (POSDIFF4) was selected as the index of Portuguese sociolinguistic proficiency for subsequent analyses.

At this point in the analysis of Portuguese proficiency for the first subsample (which received the oral tests and the multiple choice discourse test), the number of cases included was reduced. Students who had resided in Toronto for less than 3.0 years, or whose scores on particular variables were extreme (outliers) were eliminated. Thus, statistics in the following tables differ somewhat from those in Tables 4 to 7, which are based on the total sample.

Descriptive statistics for all selected Portuguese variables for this reduced sample are presented in Table 8, p. 59, while Table 9, p. 60, presents the correlations among these variables as well as with students' self-rating of their French proficiency. It is clear that there is a considerable degree of inter-relationship among the Portuguese proficiency variables. The total score on the IDEA test (POGITOTL, which includes grammar, vocabulary and phonology items) correlates significantly with all other Portuguese proficiency variables while the global rating of students' oral discourse skills (PODGLOBL) correlates significantly with all variables except phonology (POGOPHON) and students' oral and written self-ratings (PORTORAL, PORTWRIT). It is interesting to note that the highest correlation ($r = .45$, $p < .001$) for students' self-rating of written Portuguese skills (PORTWRIT) is with Portuguese written multiple-choice discourse (PWDMULTC). Self-rating of oral Portuguese skills (PORTORAL) fails to correlate significantly with written discourse but it does correlate significantly with students' oral grammar skills (POGITOTL, $r = .42$, $p < .001$). These relationships suggest that the self-ratings of Portuguese profiency are assessing similar dimensions to some of the more objective tests. The self-rating of French proficiency correlates significantly only with written discourse multiple-choice (PWDMULTC) and both Portuguese self-rating measures, although the correlation with written Portuguese self-rating is considerably higher than with oral Portuguese self-rating (.62 vs .34). This pattern of relationships suggests that the Portuguese written self-rating and French self-rating measures as well as the written Portuguese discourse measure are all assessing, to some extent, a general academic skills dimension.

The Varimax rotated solution of a principal components factor analysis of selected Portuguese proficiency measures is shown in Table 10, p. 61. Two factors emerged, defined respectively by the assessment measures and the self-ratings. Portuguese oral grammar (POGIGRAM) and discourse multiple-choice (PWDMULTC) had split loadings on the two factors.

To what extent do the data presented above support hypotheses that have been elaborated concerning the nature of language proficiency? The original theoretical framework for the study, which was based on the Canale/Swain communicative competence framework as well as the context-embedded/context-reduced distinction proposed by Cummins (1984), postulated distinctions between grammatical, discourse and sociolinguistic competence in both oral and written modalities. However, on the basis of the French immersion proficiency study this framework was revised to take account of the dynamic nature of language proficiency as it develops in different classroom and interactional contexts. Thus, empirical distinctions among proficiency dimensions can be taken as evidence for particular conceptual distinctions, but failure to find hypothesized empirical distinctions cannot be taken as evidence that the conceptual distinctions are invalid since the interactional conditions for differentiation of specific components of proficiency may not have been present. Thus, in the immersion study the finding of an academic written French factor was interpreted in light of the fact that the experience of French among immersion students has been exclusively in a classroom context with relatively little opportunity for differentiation of components of proficiency.

The present data suggest that for this grade 7 Portuguese minority student sample, variation in Portuguese proficiency can be described in terms of a global proficiency dimension supplemented by aspects of proficiency that appear more related to academic competence in Portuguese. There is little evidence that grammatical, discourse and sociolinguistic components of proficiency have become clearly differentiated in the oral modality. Since only one of the measures assesses proficiency in a context-reduced (written) situation, it is not possible to draw firm conclusions regarding relationships among proficiency components in context-reduced situations.

English proficiency. Descriptive statistics for the English proficiency measures are presented in Table 11, p. 62. Correlations among the English oral grammar measures are presented in Table 12, p. 63. It is clear that few of the indices of English grammatical competence relate significantly to each other. Although ratings of syntactic sophistication (EORLSYNT) and pronunciation (EPRONUNC) related to several of the other variables, these were not used in subsequent analyses for several reasons: first they were derived from the same language sample (story telling) as the English oral discourse variables and it was felt that this might result in spurious relationships between grammar and discourse indices; second, there was relatively low variability and near ceiling effects in both these variables, indicating that most students were being rated as essentially native speakers of English. Since no one measure of English grammatical proficiency could be considered representative, three measures were chosen for subsequent analysis (EPRPERAT -preposition error rate; EPRPTYTO - preposition type/token ratio; ECUNITSZ - average size of C-unit).

Correlations among the English oral discourse measures (see Table 13, p. 64) were all highly significant and most also related significantly to the written discourse multiple-choice measure, although at a considerably lower level. The global rating of oral discourse proficiency (EODGLOBL) and the mean of four indices of discourse competence (EODDETL4) were chosen for subsequent analyses.

Table 14, p. 65, shows the correlations for the six sociolinguistic indices that were scored. There is a considerable degree of inter-correlation among the indices, although somewhat less than in the case of English oral discourse indices. Two indices were chosen for further analyses. These were appropriateness ratings averaged over high and low situations (EOSAVAPP) and the difference between high and low situations in use of markers of formality (EOSDIFF). Two indices were chosen because they correlated at a relatively low level with each other ($r = .29$) and thus appeared to reflect only partially overlapping aspects of sociolinguistic competence.

As with Portuguese proficiency measures, recent arrivals were eliminated; there were no outliers on the English proficiency measures. Descriptive statistics for selected English oral grammar, discourse and sociolinguistic variables, discourse multiple-choice and self-rated proficiency are presented in Table 15, p. 66, and correlations among them are shown in Table 16, p. 67, for the reduced sample.

It is immediately obvious that there are considerably fewer intercorrelations among the English variables than among the Portuguese variables that assessed similar proficiency traits. Apart from the large correlations between different indices of English oral discourse proficiency, only two correlations involving the test variables attained statistical significance. The English multiple-choice written discourse measure correlates significantly with the English oral self-rating measure while the global oral discourse rating correlates significantly with the French self-rating. The lower correlations in this sample compared to Table 13 (.23 vs .28; .27 vs .31) between the two oral English discourse measures and the written discourse measure is likely due to the elimination of recently arrived students from the sample whose scores are analysed in Table 16. The same pattern also occurs for many of the other English variables.

A major reason for the lack of relationships among the English oral proficiency measures appears to be that virtually all students have attained a high degree of ease in expressing themselves in context-embedded situations. Thus, variation that does exist is unrelated to broader dimensions of proficiency in English. The high levels of confidence that students have in English is indicated by their mean self-ratings of 4.6 (out of 5) for English oral and 4.4 for English written compared to 3.8 and 3.1 for the equivalent Portuguese self-ratings. There was little point in carrying out a factor analysis on the English data due to the minimal correlations among different indices of proficiency.

One would still expect, however, that written English skills would relate strongly together by virtue of their relationship to overall academic development. Direct examination of this hypothesis was not possible due to the fact that the written grammar and sociolinguistics measures were administered to different samples. However, the question can be examined indirectly through the cross-lingual relationships for written and oral measures. If we find strong cross-lingual relationships for the English and Portuguese written measures then it suggests that all these measures are assessing a general academic dimension that exists both within and across languages. These issues are examined in the analyses of cross-lingual relationships between English and Portuguese.

3:3. Relationships Among Proficiency Variables Across Languages

The relationships among the self-ratings of proficiency are presented in Tables 17 and 18, p. 68. Table 17 presents descriptive statistics and correlations for students who were in the orally tested group, except recent arrivals, while Table 18 presents data for the entire sample with the exception of recent arrivals. As would be expected from the previous analyses, ratings of English oral proficiency tended to show low correlations with other variables except English written proficiency, while ratings of Portuguese written proficiency, French proficiency and English written proficiency all tended to relate significantly to each other. It is reasonable to interpret these relationships as reflecting, at least to some extent, an academic dimension of proficiency that crosses language boundaries.

Final descriptive statistics and intercorrelations between English and Portuguese measures for the main orally tested sample are presented in Tables 19 and 20, pp. 69 and 70. Students who were outliers on Portuguese variables or who were recent arrivals are excluded from measures in both languages. The number of variables was further reduced in order to permit exploratory factor analysis of the data. Because of their minimal relationships with other variables, English grammar variables were not included in subsequent analyses.

A distinct pattern of relationships across languages appears. First is the relatively strong correlation between English and Portuguese written (multiple-choice) discourse measures ($r = .54$, $p < .001$) and the somewhat weaker correlations between the written discourse measures in each language (PWDMULTC, EWDMULTC) and the global rating of oral discourse in the other language (PODGLOBL, EODGLOBL). It is interesting to note that the correlation between English written discourse and Portuguese global rating is higher than the correlation between English written discourse and English global rating (.44 vs .21).

The pattern of cross-language relationships can be further examined in the Varimax rotation of the principal components factor analysis shown in Table 21, p. 71. Four factors with eigenvalues greater than one emerged. Factor 1 has loadings from all Portuguese proficiency measures but its highest loading is from English written discourse (EWDMULTC). Factor 2 is defined by the Portuguese self-rating measures with smaller loadings from Portuguese grammar and vocabulary (POGIGRAM, POGIVOCA). Factor 3 has loadings from all the oral discourse and sociolinguistic indices in Portuguese and English (PODGLOBL, POSDIFF4, EODGLOBL, EOSDIFF) as well as Portuguese vocabulary. Finally, Factor 4 is defined by the English self-ratings with loadings from English and (to a small extent) Portuguese written discourse.

Two other sets of analyses were conducted to examine the pattern of relationships between components of proficiency across languages. Tables 22 through 25, pp. 72-74, present data from the subsamples that received the written grammar and sociolinguistic measures in each language. Descriptive statistics for each variable are shown in Tables 22 and 24, while the correlation matrices are in Tables 23 and 25.

The correlations in Table 23 show that English and Portuguese written grammar measures correlate .51 ($p < .01$) but neither is related to any of the self-rating variables. This is in contrast to the significant correlations between the written discourse measures and several of the self-rating indices.

The correlations presented in Table 25 show strong cross-lingual relationships in the case of one of the Portuguese sociolinguistic measures, namely the difference between the letter and note in use of formality markers (PDIFFORM). The two other Portuguese measures which involved impressionistic judgements of appropriateness were unrelated both to this index of sociolinguistic proficiency and the English sociolinguistic indices. Correlations with self-ratings of proficiency failed to attain significance for any of the written sociolinguistic measures. The written sociolinguistic relationships should be interpreted cautiously due to the small numbers of subjects involved in these analyses.

Conclusions: The nature of language proficiency and its cross-lingual dimensions. The data reinforce the conclusion of the Year 1/2 French immersion study that the relationships among components of language proficiency cannot be considered outside of particular acquisition contexts. There is evidence of a considerably tighter set of interrelationships for the Portuguese proficiency measures as compared to the English measures. This difference is likely to be related both to reduced variability for English as compared to Portuguese oral language skills and to the related fact that the interaction/acquisition contexts differ considerably for Portuguese and English for these students. English is the language of peer interaction, television/radio, and school whereas the home and Portuguese heritage language classes are the major contexts for use of Portuguese. Thus, the intuitive knowledge that students require for appropriate use of English in context-embedded situations is probably further developed than is the case for their knowledge of Portuguese. This appears to be particularly so for oral

English grammar variables which showed virtually no significant relationships with other English or Portuguese proficiency variables.

Also consistent with the French immersion results is the emergence of evidence for a context-reduced language proficiency dimension. In the immersion study this emerged in significant relationships across different components of written French proficiency, whereas in the present study the evidence is derived from significant correlations across languages, particularly for the written discourse measures. Since written grammatical, discourse, and sociolinguistic measures could not be given to the entire sample, it is not possible to replicate the immersion data directly within each language. However, it is clear that the data are consistent with the linguistic interdependence hypothesis (Cummins 1979, 1984).

The evidence from the factor analysis also suggests that cross-lingual relationships obtain in other aspects of proficiency in addition to context-reduced (academic) proficiency. Specifically, some cross-lingual relationship was observed for oral discourse and sociolinguistic variables (Factor 3), a finding that is consistent with results obtained in studies by Snow (1983) and Cummins et al (1984).

Another finding of interest is the fact that Portuguese vocabulary spreads its loading across three factors, suggesting that vocabulary knowledge is involved in most aspects of language proficiency.

In summary, the data do not support any absolute distinction between grammatical, discourse and sociolinguistic competence apart from contexts of acquisition; however, there is strong evidence for interdependence of academic skills across languages, and some evidence that cross-lingual relationships may also obtain for aspects of context-embedded skills.

3:4 Predictors of Bilingual Proficiency Development

Descriptive statistics for the Use and Attitude Questionnaire variables are presented in Appendix D, pp. 119-153, together with the questionnaire items themselves. The general trends that emerge from the questionnaire data are described in section 3:1 above. On the basis of both conceptual and empirical considerations, a subset of use and attitude variables was created to use as predictor variables in the regression analyses. The primary criterion for grouping variables together was conceptual, namely the extent to which the clustering was theoretically interpretable with respect to potential influences on bilingual proficiency development. Empirical considerations were taken into account insofar as individual variables with low correlations with the dependent variables were generally excluded from composite variables. The predictor variables used in the regression analyses are presented in Tables 26 (Portuguese) and 27 (English), pp. 75-79. Descriptive statistics and correlations of these variables with the dependent variables, controlling for age of arrival (AOA) are presented in Appendix E, pp. 154-158.

For all the dependent variables, AOA (age of arrival in Toronto) was entered first into the equation in order both to assess its effects independently of the other predictor variables and to remove the effects of length of exposure to English (or relative lack of intensive exposure to Portuguese) from the relationships between the dependent variables and the other predictors. Only the summary tables for each dependent variable in the different samples will be presented but significant partial correlations (controlling for AOA) will also be noted for each variable in order to present a fuller picture of the interrelationships among predictor and dependent variables. Pairwise deletion of missing data was used in all regressions. The proportion of variance accounted for in the equations is presented in terms of adjusted R square. The sample on which the regressions are based consisted of students whose length of residence in Canada was three years or greater.[4]

Portuguese, English and French self-rated proficiency. The summary statistics for the multiple regressions involving the Portuguese, English and French proficiency self-ratings are presented in Table 28, p. 80. The number of cases ranged between 150 and 170 with the exception of father's job (FSJOB) which had an N of 121. For Portuguese Oral Self-Rating (PORTORAL) four variables enter the equation after age of arrival (AOA) and account for 34% of the variance (adjusted R square). Students' acceptance of or liking for Portuguese (ACCPTP2) and their knowledge and pride in Portuguese culture and achievements (KANDPP) are attitudinal variables that relate to students' self-rated oral Portuguese proficiency, while students' use of Portuguese media (MEDIAP1) and their exposure to Portuguese in the home (PFAMUSE4) are also related. Sixteen variables also showed significant partial correlations with Portuguese Oral Self-Rating (PORTORAL) after age of arrival (AOA) had been entered into the equation.

Two of the same variables, namely, use of Portuguese media and knowledge and pride in Portuguese culture and achievements (MEDIAP1, KANDPP) are also represented in the equation for Portuguese Written Self-Rating (PORTWRIT). In addition, students' acceptance of and liking for French (ACCPTF2) enters the equation with a total variance accounted for of 42%. An additional sixteen variables showed significant partial correlations with Portuguese Written Self-Rating.

Considerably less variance was accounted for in the English proficiency self-rating variables. This is likely to be due, at least in part, to the ceiling effects for English self-ratings (for each of the four skills, more than 60% of the sample rated themselves in the top category ("extremely easily") compared to less than 30% in the top category for each of the Portuguese skills (see Appendix D). For English Oral Self-Rating (ENGLORAL) three variables entered after age of arrival (AOA) and accounted for 21% of the variance. The most strongly related variable was students' acceptance of or liking for English (ACCPTE2) followed by the extent to which English was used in talking to siblings and students' acceptance of French. Four other variables showed significant partial correlations with ENGLORAL.

Only 15% of the variance in English Written Self-Rating (ENGLWRIT) was accounted for by the predictor variables with students' acceptance of and liking for French (ACCPTF2) the only variables to enter the equation after age of arrival. Three other variables showed significant partial correlations with the dependent variable. It is clear that students' ratings of their Portuguese oral and written proficiency are related to a considerably broader range of predictor variables than is the case for ratings of English proficiency.

The predictor variables accounted for 47% of the variance in French self-rating with students' acceptance and liking of French, father's job, and importance of English for family communication (ICOMFAME) entering the equation. The relationship of father's job to French self-rating suggests that motivation to learn French may be related to socio-economic factors. Eleven variables, several reflecting positive attitudes towards Portuguese and English-Canadian cultures, had partial correlations that reached a level of statistical significance.

Portuguese proficiency. The regression analyses for the Portuguese oral and written proficiency measures are summarized in Table 29, p. 81. Five variables explain 41% of the variance for the total score on the IDEA test which comprises grammar, vocabulary, and some phonology items (POGITOTL). Visits to Portugal (VISITS) and formal study of the language (CLASSA) appear to be especially important predictors of Portuguese grammatical proficiency. Use of various Portuguese media also appears to be moderately related (partial $r = .39$), although it does not enter the regression equation because of its own relationship with visits to Portugal (VISITS). In addition to these variables, attendance at Portuguese mass and knowledge and pride in Portuguese culture and achievements (MASS, KANDPP) had significant ($p < .05$) partial correlations with Portuguese proficiency (POGITOTL). The entry of Portuguese vitality (VITALP2) into the regression equation is somewhat misleading since its initial partial correlation with Portuguese proficiency (POGITOTL) is negative and low (-.19); it enters as a result of its interaction with other predictors (visits to Portugal and formal study of Portuguese (VISITS and CLASSA)). Use of Portuguese at community events (COMMEVNT) also shows a negative relation with proficiency (POGITOTL) which, in this case, indicates that use of Portuguese at community events is positively related to proficiency.

Only two variables enter the regression equation for Portuguese Oral Discourse Global Score (PODGLOBL) after age of arrival (AOA): French vitality and formal study of Portuguese (VITALF1 and CLASSA), accounting for 24% of the variance. Three additional variables had significant partial correlations: father's job, visits to Portugal, and language mixing (FSJOB, VISITS, and LMIXG).

For Portuguese Written Discourse Multiple-Choice (PWDMULTC), formal study of Portuguese (CLASSA) is again a significant predictor, accounting for 20% of the variance. Amount of TV watching per week (TVHRSP) adds a further 10% while the child and family's preferred country of residence (ORIENTPX) brings the total to 39% variance explained. The relationship of this latter variable to Portuguese Discourse Multiple-Choice (PWDMULTC) is negative, indicating that proficiency is related to a desire to remain in Canada rather than return to Portugal. Two variables related to these latter predictors showed significant ($p < .05$) positive partial correlations with Portuguese Discourse Multiple-Choice (PWDMULTC): liking Portuguese TV and knowledge and pride in Portuguese culture and achievements (LIKETVP and KANDPP).

For oral sociolinguistic proficiency (POSDIFF4), only father's job (FSJOB) enters the equation, accounting for about 10% of the variance. The relation between father's job (FSJOB) and sociolinguistic proficiency (POSDIFF4) is negative, indicating that lower socio-economic status students tend to perform better on this task. None of the partial correlations attained significance.

In summary, amount of exposure, both formal (CLASSA) and informal (e.g. VISITS, TVHRSP, MEDIAP1 and MASS), appears to play a major role in predicting different aspects of Portuguese proficiency. This is a typical finding in the case of minority languages to which students tend to be minimally exposed outside the home. However, attitudes also appear important, although to a considerably lesser extent. Knowledge and pride in Portuguese culture and achievements (KANDPP) tended to show consistent positive relationships with the dependent variables, although it did not enter the regression equations because of colinearity with other variables. A number of Attitude and Use variables that might have been expected to relate positively to Portuguese proficiency failed to do so. For example, family use of Portuguese did not predict proficiency nor, in general, did students' perceptions of parental and societal attitudes towards Portuguese.

English proficiency. There were minimal relationships between the predictors and the English oral grammar and sociolinguistic variables (see Table 30, p. 82). Relationships were somewhat greater for the oral and written discourse measures with 16% and 25% of the variance respectively explained. It is important to note that language used in talking with siblings (TALKWSIB) and attendance at Portuguese or English mass (MASS) both relate negatively to English Oral Discourse (EODDETL4) and English Written Discourse Multiple-Choice (EWDMULTC), indicating that use of Portuguese in talking with siblings and attendance at Portuguese mass are positively related to English proficiency. Three additional variables (language used by students at community events, vitality of French, and language mixing (COMMEVNT, VITALFI, LMIXG)) showed significant ($p < .05$) partial correlations with English Discourse Writen Multiple-Choice, controlling for age of arrival. The extent to which students viewed French as having high ethnolinguistic vitality (VITALFI), almost entered the regression equation after MASS but its partial correlation failed to attain the $< .05$ criterion (partial r controlling for AOA, FSJOB and MASS: .30, $p = .054$).

Written grammatical and sociolinguistic proficiency. There were few significant relationships between the predictor variables and either Portuguese or English written grammatical proficiency (see Table 31, p. 83). For Portuguese grammar after age of arrival (AOA) had been entered into the equation, three variables showed partial correlations which approached significance ($p < .10$). These were: amount of TV watched per week, language mixing, and French vitality (TVHRSP, LMIXG, and VITALF1).

The extent to which students like to use English when talking to members of their families was the only variable to enter the regression equation for English written Grammar after age of arrival. The partial correlations for two other variables approached significance ($p < .10$): father's job and importance of English for communicating within the family (FSJOB and ICOMFAME).

No significant relationships were observed for any of the three Portuguese written sociolinguistic variables (differentiation in use of formality markers between note and letter (PDIFFORM), appropriateness rating for the letter alone (PLETGLOB) and appropriateness rating for the letter and the note combined (PGLOBAV)). However, French vitality (VITALF1) does approach significance for PDIFFORM (partial $r = -.51$, $p < .07$).

For English sociolinguistic proficiency no significant relationships were observed between the predictors and either rating of appropriateness of mitigation in the letter (ELETMIT) or average of the global appropriateness ratings for letter and note (EGLOBAV). However, the predictors did account for 58% of the variance on the rewriting task (REWRITE). This task required students to rewrite short segments in each of six paragraphs so as to make the register fit the rest of the paragraph more appropriately. It may be that this task taps general writing competence in English in addition to specific sociolinguistic knowledge. Students' perception of whether people value French is negatively related to REWRITE, accounting for an additional 21% of the variance after age of arrival (AOA) has been entered. The extent to which the student reports speaking English at community events (COMMEVNT) is positively related to REWRITE but the extent to which the student likes to talk with friends in English (LIKEFRND) is negatively related to REWRITE. This variable enters the equation as a result of interactions with French vitality (VITALF1) and language used by students at community events (COMMEVNT) and thus its relation to the dependent variable is difficult to interpret.

Summary of Regression Analyses. The regression analyses show that more variance is explained for the Portuguese proficiency variables than for the English variables. This is likely due in part to the greater variance on most of the Portuguese variables than on the English variables. Use and exposure to Portuguese appeared to be more related to proficiency than were attitudes to or perceptions of the language and culture. In particular, study of the language (CLASSA), visits to Portugal (VISITS), watching Portuguese TV (TVHRSP), and activities such as reading Portuguese books, listening to Portuguese radio and writing letters in Portuguese (MEDIAP1) showed consistent relationships to proficiency. Of the attitudinal variables, knowledge and pride in Portuguese culture and achievements (KANDPP) showed the most consistent relationship to proficiency.

Few strong trends emerged for the prediction of English proficiency. Father's job classification (FSJOB) was related to English Discourse Written Multiple-Choice (EWDMULTC), the most academic of the English measures. However, indices of exposure and use tended to relate in the opposite direction to what might have been expected. For example, language used with siblings and language of mass (TALKWSIB and MASS) both relate negatively to the two English discourse measures (EODDETL4 and EWDMULTC), indicating that use of Portuguese in talking with siblings and attendance at Portuguese Mass are positively related to English proficiency. This trend suggests that positive attitudes towards Portuguese maintenance and actual use of Portuguese in community and home are in no way detrimental to students' English proficiency.

4. COMPARISON OF PORTUGUESE PROFICIENCY WITH AZOREAN NATIVE SPEAKERS

In order to further investigate the extent to which the social and educational context influenced the development of bilingual proficiency, data were collected from native Portuguese students in San Miguel island in the Azores. Fifty-six percent of the grade 7 students tested in Toronto were of Azorean background and thus the comparison with Azorean rather than mainland Portuguese students is appropriate.

Collection of the Azorean data was made possible through a joint collaboration between OISE and both the Portuguese Secretary of State for Emigration and the Secretary of State for Emigration of the Regional Government of the Azores. Testing was carried out over a period of three weeks (April and May 1984). Oral interviews were conducted by a male native speaker of Portuguese who had also been involved in the Toronto data collection. The Department of Education in San Miguel selected the Junior High School (Escola Preparatoria) for test administration.

The following measures were administered to a total of 69 grade 6 students in three different classes:

(1) all the Portuguese versions of the written tests administered in Toronto to grade 7 students i.e., Grammar Multiple Choice, Discourse Multiple Choice, Sociolinguistic Compositions;

(2) oral interview, story retelling (from the IDEA test), picture description (from the IDEA test), oral sociolinguistic test.

Only the two components of the IDEA test mentioned above were administered since the other items were considered too easy for the Azorean students. The written and oral measures were scored in the same way as for the Toronto students. The one-way ANOVA analyses comparing the Toronto and Azorean performance on these measures is presented in Table 32, p. 84.

There are highly significant differences between the Toronto and Azorean students on most measures of Portuguese proficiency. Differences are most apparent on the written grammar measure and least apparent on the written sociolinguistic measures with discourse measures occupying an intermediate position. Major differences do appear on the written discourse multiple choice measure but are less obvious on the oral discourse measures.

These data parallel the differences found in the Year 1 study between French immersion and native francophone students in that differences in grammar were most salient with less major differences in sociolinguistic and discourse proficiencies. The data are also consistent with the findings of the regression analyses which indicated that indices reflecting exposure to and use of Portuguese were more clearly related to grammatical proficiency in Portuguese (at least oral grammar) than they were to discourse and sociolinguistic proficiencies.

In conclusion, the large differences between the Toronto and Azorean students show how formidable is the task of maintaining first language proficiency in a minority context. Despite the fact that the Toronto sample also consists of native speakers of Portuguese and many are quite fluent in oral Portuguese in context-embedded situations, their explicit knowledge of the formal structure of the language appears relatively limited in comparison to the Azorean native speakers. The Toronto students, however, appear to be extremely comfortable in English, as indicated by their self-ratings and scores on the English context-embedded measures. There is thus no evidence from our data of any linguistic disadvantage in English, although comparison data with English native speakers are not available. The strong relationship between attendance at Portuguese language classes and the performance of the Toronto students suggests that more intensive exposure to Portuguese in an academic context could have a significant impact on bridging the gap between their Portuguese proficiency and that of native Portuguese-educated students.

Footnotes

1 The authors would like to thank Ellen Bouchard-Ryan for her help in conceptualizing the content of the questionnaires.

2 Notions of urban vs. rural were agreed by the scorers based not on the fact that the place of birth was a city, town or village, but on the amount of industrialization in or near each place. According to this judgement, and to give a Canadian example, a place like Oshawa (Ontario) would be considered as an urban setting, while a place like Lakefield (Ontario) would be considered a rural setting. This notion is particularly relevant as it applies to places in Portugal. It should be considered that, in this case, the capitals of some Portuguese provinces are in fact rural settings while a small town located in the industrialized regions of Lisbon or Oporto might be part of an urban setting. Concerning the nine islands of the Azores which are predominantly rural, only the three main towns: Ponta Delgada (in São Miguel), Angra do Heroismo (in Terceira) and Horta (in Faial) were considered as urban settings even though none of them could be considered as highly industrialized centres.

3 In analysing the relationships of Portuguese grammar variables with other indices of Portuguese proficiency, it was found that most measures of Portuguese grammar tended to show the same pattern of correlations with discourse and sociolinguistic variables as the variables derived from the IDEA test. For example, grammar error rate (POGERRAT) and the rate excluding dialect 'errors' (POGEREXD) correlated significantly with PODDETL3 (oral Portuguese discourse indices assessing logic, anaphora, and time), PODGLOBL (global rating of Portuguese oral discourse), PWDMULTC (written discourse multiple-choice) and Age on Arrival (AOA) as do all the variables derived from the IDEA test with the exception of Phonology. None of these variables correlates significantly with any of the English grammar variables, as discussed below.

4 The N for the regression analyses is less than for correlational analyses as a result of a smaller number of cases for some predictor variables.

References

Barker, G.C. Social functions of language in a Mexican-American community. **Acta Americana,** 1947, 5, 185-202.

Berry, J.W., Kalin, R., Taylor, D.M. **Multiculturalism and Ethnic Attitudes in Canada.** Ottawa: Ministry of Supply and Services, 1977.

Bourhis, R.Y. The charter of the French language and cross cultural communication in Montreal. In Bourhis, R.Y. (ed.), **Conflict and Language Planning in Quebec.** Clevedon, England: Multilingual Matters, 1986.

Bourhis, R.Y, Giles, H., Rosenthal, D. Notes on the construction of a 'subjective vitality questionnaire' for ethnolinguistic groups. **Journal of Multilingual and Multicultural Development,** 1981, 2, 145-155.

Brody, E.B. (ed.) **Minority Group Adolescents in the United States.** Baltimore: The Williams and Witkin Co., 1968.

Carranza, M.A. Attitudinal research on Hispanic language varieties. In E.B. Ryan and H. Giles, (eds.), **Attitudes Towards Language Variation: Social and Applied Contexts.** London: Edward Arnold, 1982.

Cummins, J. Linguistic interdependence and the educational development of bilingual children. **Review of Educational Research,** 1979, 49, 222-251.

Cummins, J. **Bilingualism and Special Education: Issues in Assessment and Pedagogy.** Clevedon, England: Multilingual Matters, 1984.

Cummins, J., Swain, M., Nakajima, K., Handscombe, J., Green, D., Tran, C. Linguistic interdependence among Japanese and Vietnamese immigrant students. In C. Rivera (ed.), **Communicative Competence Approaches to Language Proficiency Assessment: Research and Application.** Clevedon, England: Multilingual Matters, 1984.

Day, R. Children's attitudes toward language. In E.B. Ryan and H. Giles, (eds.), **Attitudes Towards Language Variation: Social and Applied Contexts.** London: Edward Arnold, 1982.

Derbyshire, R. Adolescent identity crises in urban Mexican Americans in East Los Angeles. In E.B. Brody (ed.), **Minority Group Adolescents in the United States.** Baltimore: The Williams and Witkin Co., 1968.

Gambell, T. The occurrence and analysis of a repertoire of situational language in grade six children. Unpublished Ph.D. dissertation. Department of Elementary Education, University of Alberta, 1978.

Gardner, R.C. and Lambert, W.E. **Attitudes and Motivation in Second Language Learning.** Rowley, Mass.: Newbury House, 1972.

IDEA Oral Language Proficiency Test, Portuguese Forms. Whittler, California: Ballard & Tigle, Inc., n.d.

Jakobovits, L. **Foreign Language Learning.** Rowley, Mass.: Newbury House, 1970.

Lambert, W.E. A social psychology of bilingualism. **Journal of Social Issues,** 1967, **23,** 91-109.

Mougeon, R., Brent Palmer, C., Bélanger, M., Cichocki, W. **Le français en situation minoritaire, Vol. 1.** Quebec City: International Center for Research on Bilingualism, 1982.

Mougeon, R., Beniak, E., Canale, M. Le problème des élèves anglo-dominants dans les écoles ontariennes de langue française: acquisition, emploi et enseignement du français. **Canadian Modern Language Review** (forthcoming).

Ryan, E.B. Language planning from an attitudinal perspective. Paper presented at the International Conference on Language Planning and Public Policy, Cancun, Mexico, 1981.

Ryan, E.B., Giles, H., Sebastian, R.J. An integrative perspective for the study of attitudes toward language variation. In E.B. Ryan and H. Giles (eds.), **Attitudes Towards Language Variation: Social and Applied Contexts.** London: Edward Arnold, 1982.

Snow, C.E. Language and literacy: Relationships during the preschool years. **Harvard Educational Review,** 1983, **53,** 165-189.

Table 1

Mean Differences Between Ratings of Standard Ethnic Groups and Average Ratings in Total Sample of Respondents*

Scale	English	French	Chinese	Portuguese	Italian
1. Hardworking	.06	.00	.30	.28	.23
2. Important	.69	.59	-.42	-.46	-.16
3. Canadian	.90	.78	-.71	-.88	-.43
4. Clean	.66	.48	-.22	-.20	-.33
5. Similar to me	1.45	1.08	-1.08	-.19	-.27
6. Likeable	.59	.51	-.21	-.12	-.19
7. Stick together as a group	-.81	-.04	.40	.39	.42
8. Wealthy	.73	.13	-.05	-.58	.29
9. Interesting	.28	.48	-.18	-.20	-.28
10. Well known to me	1.68	1.16	-.94	-.05	0

*Based on Berry, Kalin and Taylor (1977:287)

Table 2

Distribution of Portuguese Sample Across Schools

School	n	% of Gr. 7 Portuguese Heritage Lg. students in sample	Overall % of Portuguese-background students in school
1	22	40.7%	60%
2	35	62.5%	75%
3	25	62.5%	90%
4	20	40.0%	75%
5	24	33.0%	90%
6	28	62.2%	85%
7	37	46.3%	90%

Table 3

Number of Students in the Sample who Completed Proficiency Tests

	Grammar written	Discourse written	Sociolinguistic written	Oral interview	Sociolinguistic oral
English	55	87	50	85	85
Portuguese	47	84	47	89	87

Table 4

Number of Cases, Mean, and Standard Deviation for Portuguese Oral Proficiency and Discourse Multiple Choice Using All Available Data for Each Variable

Variable	Cases	Mean	SD
POGITOTL	77	.8339	.1063
POGIGRAM	77	.8249	.1277
POGIPHON	77	.9596	.0741
POGIVOCA	77	.7815	.1514
POGERRAT	73	.3305	.2015
POGEREXD	73	.2998	.1860
PPRPTYPS	78	2.9231	1.2970
PVRBTNSS	78	2.4231	.7120
PPRPERRP	74	.2432	.5442
PPRPTYPP	74	2.5541	1.2402
PPRPTOKP	74	3.8514	2.2976
PORTORAL	78	3.7244	.9658
PORTWRIT	78	3.1154	1.1422
FRENCH	71	2.9648	.8591
PODGLOBL	76	2.9211	1.1635
PODDETL3	76	3.9035	.8125
PODCONCL	77	1.4286	.8180
PWDMULTC	77	.4955	.1642
POSDIFF4	73	1.0107	.3324
POSAVAPP	76	4.3833	.4283
POSAPPLO	77	4.6846	.4234
POSAPPHI	76	4.0664	.6056

Table 5

Correlations between Oral Portuguese Grammar Measures Using All Available Cases for Each Analysis

	POGITOTL	POGIGRAM	POGIPHON	POGIVOCA	POGERRAT	POGEREXD
POGIGRAM	.9027**					
POGIPHON	.5480**	.3813**				
POGIVOCA	.8365**	.5367**	.4291**			
POGERRAT	-0.5486**	-0.4991**	-0.1519	-0.4986**		
POGEREXD	-0.4492**	-0.4067**	-0.1062	-0.4162**	.9431**	
PPRPTYPS	.4689**	.3902**	.2564	.4394**	-0.0970	-0.0603
PVRBTNSS	.3148*	.2850*	.0851	.2857*	-0.1836	-0.1785
PPRPERRP	-0.3106*	-0.3404*	-0.0815	-0.2010	.2836*	.2136
PPRPTYPP	-0.1410	-0.1539	-0.1712	-0.0569	.0305	-0.0493
PPRPTOKP	.0018	-0.0179	-0.0984	.0536	-0.1191	-0.1805
PORTORAL	.5114**	.4606**	.1644	.4592**	-0.3689**	-0.3217*
PORTWRIT	.4638**	.4402**	.1636	.3829**	-0.2738*	-0.2328
FRENCH	.0602	.0628	.0078	.0448	-0.1149	-0.1573

	PPRPTYPS	PVRBTNSS	PPRPERRP	PPRPTYPP	PPRPTOKP
PVRBTNSS	.2326				
PPRPERRP	-0.2926*	-0.1693			
PPRPTYPP	-0.1912	-0.0292	.2847*		
PPRPTOKP	-0.1627	.1861	.2484	.7889**	
PORTORAL	.3920**	.0985	.0193	-0.0749	-0.1169
PORTWRIT	.3066*	.0297	-0.0938	-0.1713	-0.0769
FRENCH	.0010	.1436	-0.0866	-0.1396	-0.0341

	PORTORAL	PORTWRIT
PORTWRIT	.6590**	
FRENCH	.1986	.4974**

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 6

Correlations between Oral Portuguese Discourse Measures, Discourse Multiple Choice, and Self-rated Portuguese and French Using All Available Data from the Main Sample for Each Analysis

	PODGLOBL	PODDETL3	PODCONCL	PWDMULTC	PORTORAL	PORTWRIT
PODDETL3	.4667**					
PODCONCL	.4683**	.3753**				
PWDMULTC	.5031**	.3732**	.3270*			
PORTORAL	.3074*	.3265*	.2585	.2852*		
PORTWRIT	.2549	.2103	.2461	.4807**	.6590**	
FRENCH	.0074	-0.0447	.1076	.2595	.1986	.4974**

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 7

Correlations between Measures of Oral Portuguese Sociolinguistics and Self-rated Portuguese and French Using All Available Data from the Main Sample for Each Analysis

	POSDIFF4	POSAVAPP	POSAPPLO	POSAPPHI	PORTORAL	PORTWRIT
POSAVAPP	.7768**					
POSAPPLO	.5089**	.7664**				
POSAPPHI	.7573**	.9039**	.4180**			
PORTORAL	.2814*	.3682**	.1999	.3684**		
PORTWRIT	.1520	.2171	.1309	.2267	.6590**	
FRENCH	-0.0110	-0.1794	-0.2604	-0.0630	.1986	.4974**

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 8

Number of Cases, Mean, and Standard Deviation for Selected Measures of Oral Portuguese Proficiency, and Discourse Multiple Choice, for the Sample Excluding Outliers and Students with less than 3 Years of Residence in Toronto

Variable	Cases	Mean	SD
POGITOTL	65	.8462	.0898
POGIGRAM	65	.8387	.1122
POGIPHON	65	.9607	.0772
POGIVOCA	65	.7973	.1286
PODGLOBL	65	3.0000	1.1592
PODDETL3	65	3.9436	.7420
PWDMULTC	65	.5118	.1640
POSDIFF4	61	1.0576	.2975
POSAVAPP	64	4.4516	.2951
PORTORAL	67	3.7836	.9055
PORTWRIT	67	3.1493	1.1045
FRENCH	61	2.9344	.8464

Table 9

Correlations between Selected Measures of Oral Portuguese Proficiency and Discourse Multiple Choice, Excluding Outliers and Students with less than 3 Years Residence from Each Analysis

	POGITOTL	POGIGRAM	POGIPHON	POGIVOCA	PODGLOBL	PODDETL3
POGIGRAM	.8752**					
POGIPHON	.6214**	.4299**				
POGIVOCA	.7678**	.3840**	.4397**			
PODGLOBL	.6465**	.5744**	.1660	.5497**		
PODDETL3	.3254*	.2765	.2268	.2498	.3996**	
PWDMULTC	.4565**	.4520**	.2327	.2949*	.4693**	.3401*
POSDIFF4	.4437**	.2920	.3055*	.4630**	.4830**	.2109
POSAVAPP	.3463*	.2216	.1921	.3839*	.4833**	.1897
PORTORAL	.4207**	.3672*	.1070	.3736*	.2752	.2159
PORTWRIT	.3496*	.3401*	.1049	.2567	.2050	.0985
FRENCH	.1564	.1624	.0429	.1045	.0472	-0.0227

	PWDMULTC	POSDIFF4	POSAVAPP	PORTORAL	PORTWRIT
POSDIFF4	.3342*				
POSAVAPP	.1643	.7310**			
PORTORAL	.2153	.1531	.1924		
PORTWRIT	.4534**	-0.0404	-0.0349	.6161**	
FRENCH	.3329*	.0666	-0.0796	.3416*	.6233**

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 10

Principal Components Factor Analysis of Selected Measures of Portuguese Proficiency, Using the Sample Excluding Outliers and Students with less than 3 Years Residence

	Factor 1	Factor 2	Communality
POGIGRAM	.61639	.44529	.57822
POGIPHON	.61794	.02532	.38249
POGIVOCA	.72321	.25852	.58986
PODGLOBL	.73298	.27231	.61141
PWDMULTC	.48803	.47314	.46203
POSDIFF4	.78819	-0.10468	.63219
PORTORAL	.13839	.80470	.66669
PORTWRIT	.00941	.91955	.84566
% Variance	.338	.258	

Table 11

Number of Cases, Mean, and Standard Deviation for Measures of Oral English Proficiency, Discourse Multiple Choice, and Self-rated Proficiency Using All Available Data for Each Variable

Variable	Cases	Mean	SD
EPRPERAT	76	.1724	.1343
EPRPTYTO	73	.6558	.1556
EVRBERAT	78	.0418	.0755
EVRBTYTO	78	.7243	.1621
ECUNITSZ	79	8.9426	1.5322
EORLSYNT	79	4.8228	.4085
EPRONUNC	79	4.4747	.5823
ENGLORAL	78	4.5513	.7006
ENGLWRIT	78	4.3654	.8005
FRENCH	71	2.9648	.8591
EODGLOBL	79	3.1203	.9379
EODDETL4	79	3.2089	.9189
EODSCENE	79	3.3987	1.0451
EODLOGIC	79	3.1709	.9867
EODANAPH	79	3.2152	.9463
EODCNECT	79	3.0506	1.0907
EWDMULTC	78	.6210	.1509
EOSDIFF	79	.2583	.1153
EOSDETLO	79	.2138	.1032
EOSDETHI	79	.4721	.1042
EOSAVAPP	79	4.1418	.3458
EOSAPPLO	79	4.7120	.3299
EOSAPPHI	79	3.5715	.5316

Table 12

Correlations between Oral English Grammr Measures and Self-rated Proficiency Using All Available Data for Each Analysis

	EPRPERAT	EPRPTYTO	EVRBERAT	EVRBTYTO	ECUNITSZ	EORLSYNT
EPRPTYTO	.0851					
EVRBERAT	.0081	-0.0742				
EVRBTYTO	.0065	.2040	-0.1670			
ECUNITSZ	.1373	-0.1147	-0.0718	.0576		
EORLSYNT	-0.2347	-0.1793	-0.4488**	-0.1545	.2214	
EPRONUNC	-0.0929	-0.2019	-0.2865*	-0.1521	.2345	.3985**
ENGLORAL	.0435	-0.0105	-0.3702**	-0.0540	.0614	.3551**
ENGLWRIT	.0229	.0211	-0.2750*	.1263	-0.0539	.1926
FRENCH	.1652	.0686	-0.0153	-0.0160	-0.1988	-0.0606

	EPRONUNC	ENGLORAL	ENGLWRIT
ENGLORAL	.2407		
ENGLWRIT	.0569	.5104**	
FRENCH	.0356	.1231	.3382*

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 13

Correlations between Measures of English Oral Discourse, Discourse Multiple Choice, and Self-rated Proficiency Using All Available Data for Each Analysis

	EODGLOBL	EODDETL4	EODSCENE	EODLOGIC	EODANAPH	EODCNECT
EODDETL4	.9300**					
EODSCENE	.8203**	.8976**				
EODLOGIC	.8815**	.9066**	.7661**			
EODANAPH	.8734**	.9125**	.7548**	.7840**		
EODCNECT	.7929**	.8983**	.7187**	.7364**	.7750**	
EWDMULTC	.2826*	.3125*	.2509	.2867*	.3818**	.2209
ENGLORAL	.1181	.0580	.0064	.1083	.1047	.0003
ENGLWRIT	.0306	.0114	-0.0650	-0.0101	.0603	.0571
FRENCH	.2688	.1541	.1082	.1658	.1541	.1312

	EWDMULTC	ENGLORAL	ENGLWRIT
ENGLORAL	.4224**		
ENGLWRIT	.3413*	.5104**	
FRENCH	-0.0212	.1231	.3382*

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 14

Correlations between Oral English Sociolinguistic Measures, and Self-rated Proficiency Using All Available Data for Each Analysis

	EOSDIFF	EOSDETLO	EOSDETHI	EOSAVAPP	EOSAPPLO	EOSAPPHI
EOSDETLO	-0.5407**					
EOSDETHI	.5676**	.3803**				
EOSAVAPP	.2905*	.1076	.4382**			
EOSAPPLO	.2266	-0.2734*	-0.0117	.6673**		
EOSAPPHI	.2373	.3097*	.5773**	.8868**	.2474	
ENGLORAL	.1594	-0.0587	.1007	.2449	.1045	.2537
ENGLWRIT	.0174	-0.0433	-0.0374	.2531	.1348	.2458
FRENCH	-0.2392	.0710	-0.1830	.0381	.0257	.0319

	ENGLORAL	ENGLWRIT
ENGLWRIT	.5104**	
FRENCH	.1231	.3382*

* - SIGNIF. LE .01

** - SIGNIF. LE .001

Table 15

Number of Cases, Mean, and Standard Deviation for Selected Measures of English Oral Proficiency and Discourse Multiple Choice for the Sample Excluding Students with less than 3 Years of Residence in Toronto

Variable	Cases	Mean	SD
EPRPERAT	69	.1758	.1362
EPRPTYTO	66	.6518	.1570
ECUNITSZ	72	8.9952	1.5839
EODGLOBL	72	3.1389	.9465
EODDETL4	72	3.2257	.9252
EWDMULTC	71	.6325	.1462
EOSDIFF	72	.2634	.1159
EOSAVAPP	72	4.1505	.3413
ENGLORAL	72	4.6319	.6109
ENGLWRIT	72	4.4097	.7887
FRENCH	65	2.9462	.8309

Table 16

Correlations between Selected Measures of English Oral Proficiency and Discourse Multiple Choice for the Sample Excluding Students with less than 3 Years Residence in Toronto

	EPRPERAT	EPRPTYTO	ECUNITSZ	EODGLOBL	EODDETL4	EWDMULTC
EPRPTYTO	.0446					
ECUNITSZ	.1414	-0.1002				
EODGLOBL	.0289	-0.2671	.1639			
EODDETL4	.0869	-0.2600	.2248	.9277**		
EWDMULTC	.0396	-0.0701	.1480	.2282	.2669	
EOSDIFF	-0.0114	-0.0230	.0793	.1732	.2148	.0521
EOSAVAPP	-0.0248	-0.1211	.0227	.1055	.0979	.1762
ENGLORAL	-0.0196	-0.0306	-0.0118	.0105	-0.0488	.3039*
ENGLWRIT	-0.0282	.0125	-0.0997	-0.0396	-0.0561	.2742
FRENCH	.1523	.0399	-0.2097	.2990*	.1815	.0108

	EOSDIFF	EOSAVAPP	ENGLORAL	ENGLWRIT
EOSAVAPP	.2482			
ENGLORAL	.0754	.1320		
ENGLWRIT	-0.0162	.1916	.4636**	
FRENCH	-0.1867	.0597	.2413	.3332*

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 17

Descriptive Statistics and Correlations for Self-rated Proficiency for the Orally Tested Sample, Excluding Students with less than 3 Years of Residence in Toronto

	PORTORAL	PORTWRIT	ENGLORAL	ENGLWRIT	Cases	Mean	SD
PORTORAL					72	3.71	.9448
PORTWRIT	.6337**				72	3.06	1.1131
ENGLORAL	.1347	.1534			72	4.63	.6109
ENGLWRIT	.1437	.3434*	.4636**		72	4.41	.7887
FRENCH	.2732	.5696**	.2413	.3332*	65	2.95	.8309

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 18

Descriptive Statistics and Correlations for Self-rated Proficiency for the Full Sample, Excluding Students with less than 3 Years of Residence in Toronto

	PORTORAL	PORTWRIT	ENGLORAL	ENGLWRIT	Cases	Mean	SD
PORTORAL					170	3.74	.9377
PORTWRIT	.6684**				170	3.04	1.1571
ENGLORAL	.1868*	.2013*			168	4.67	.5868
ENGLWRIT	.2419**	.3324**	.5324**		169	4.48	.7713
FRENCH	.3110**	.4409**	.3493**	.4133**	147	2.88	.7929

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 19

Number of Cases, Mean, and Standard Deviation for Selected Measures of Portuguese and English Proficiency, Excluding Outliers on Portuguese Variables and Recent Arrivals from Measures in Both Languages

VARIABLE	Cases	Mean	SD
POGIGRAM	65	.8387	.1122
POGIPHON	65	.9607	.0772
POGIVOCA	65	.7973	.1286
PODGLOBL	65	3.0000	1.1592
PWDMULTC	65	.5118	.1640
POSDIFF4	61	1.0576	.2975
PORTORAL	67	3.7836	.9055
PORTWRIT	67	3.1493	1.1045
EODGLOBL	67	3.1791	.9322
EWDMULTC	66	.6349	.1500
EOSDIFF	67	.2736	.1128
ENGLORAL	67	4.6194	.6282
ENGLWRIT	67	4.3955	.8097

Table 20

Correlations between Selected Measures of Portuguese and English Proficiency, for the Sample Excluding Outliers on Portuguese Variables and Recent Arrivals

	POGIGRAM	POGIPHON	POGIVOCA	PODGLOBL	PWDMULTC	POSDIFF4
POGIPHON	.4299**					
POGIVOCA	.3840**	.4397**				
PODGLOBL	.5744**	.1660	.5497**			
PWDMULTC	.4520**	.2327	.2949*	.4693**		
POSDIFF4	.2920	.3055*	.4630**	.4830**	.3342*	
PORTORAL	.3672*	.1070	.3736*	.2752	.2153	.1531
PORTWRIT	.3401*	.1049	.2567	.2050	.4534**	-0.0404
EODGLOBL	.1859	.1649	.2461	.2191	.2941*	.3587*
EWDMULTC	.2588	.2218	.2485	.4382**	.5425**	.1986
EOSDIFF	.0316	-0.0526	.0897	.1717	.0232	.1947
ENGLORAL	-0.1426	-0.1072	-0.0775	.0425	.1884	.0062
ENGLWRIT	-0.1072	-0.0738	.1726	.0827	.1752	-0.0416

	PORTORAL	PORTWRIT	EODGLOBL	EWDMULTC	EOSDIFF	ENGLORAL
PORTWRIT	.6161**					
EODGLOBL	.1588	.1650				
EWDMULTC	.0274	.0293	.2145			
EOSDIFF	-0.1976	-0.1822	.1430	.0272		
ENGLORAL	.1527	.1814	.0276	.3086*	.1038	
ENGLWRIT	.1909	.3820**	-0.0100	.2985*	.0103	.4643**

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 21

Principal Components Factor Analysis of Selected Measures of Portuguese and English Proficiency for the Sample Excluding Recent Arrivals and Outliers on Portuguese Variables

	Factor 1	Factor 2	Factor 3	Factor 4	Communality
POGIGRAM	.65670	.35767	.20667	-0.23449	.65688
POGIPHON	.63787	.08047	.06481	-0.28381	.49811
POGIVOCA	.43639	.37614	.49260	-0.08244	.58137
PODGLOBL	.56473	.18588	.51335	.08116	.62360
PWDMULTC	.68076	.19454	.17005	.33281	.64096
POSDIFF4	.34848	.02113	.69636	-0.09729	.61626
PORTORAL	.09766	.86359	.10972	.06701	.77186
PORTWRIT	.15739	.84216	-0.05834	.26265	.80639
EODGLOBL	.16034	.14499	.55089	.03611	.35151
EWDMULTC	.71921	-0.21745	.07392	.51369	.83389
EOSDIFF	-0.20632	-0.29725	.71125	.13591	.65526
ENGLORAL	-0.03975	.03749	.04851	.81634	.67174
ENGLWRIT	-0.00426	.26723	-0.00943	.76294	.65359
% Variance:	.196	.156	.146	.144	

Table 22

Number of Cases, Mean, and Standard Deviation for Measures of Portuguese and English Written Grammar and Self-rated Proficiency for the Sample Excluding Students with less than 3 Years of Residence in Toronto

VARIABLE	CASES	MEAN	STD DEV
PGRAMMC	41	.4793	.1805
PORTORAL	45	3.7778	.9508
PORTWRIT	45	3.0667	1.1851
EGRAMMC	45	.4542	.1478
ENGLORAL	45	4.7111	.5789
ENGLWRIT	45	4.5667	.7508
FRENCH	41	2.8293	.7835

Table 23

Correlations between Portuguese and English Written Grammar and Self-rated Proficiency for the Sample Excluding Students with less than 3 Years Residence in Toronto

	PGRAMMC	PORTORAL	PORTWRIT	EGRAMMC	ENGLORAL	ENGLWRIT
PORTORAL	.0440					
PORTWRIT	.2490	.6992**				
EGRAMMC	.5077**	-0.0550	.1037			
ENGLORAL	.1929	.2317	.1612	.0178		
ENGLWRIT	.0407	.3078	.2759	.0793	.5421**	
FRENCH	.2995	.3770*	.3090	.1439	.4103*	.4350*

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 24

Number of Cases, Mean, and Standard Deviation for Measures of Portuguese and English Written Sociolinguistics and Self-rated Proficiency, for the Sample Excluding Students with less than 3 Years Residence in Toronto

VARIABLE	Cases	Mean	SD
PGLOBAV	21	3.7381	1.0077
PLETGLOB	24	3.2500	1.2247
PDIFFORM	18	1.1991	.3781
PORTORAL	41	3.7561	1.0254
PORTWRIT	41	3.0854	1.2037
EGLOBAV	30	3.3083	.6781
ELETMIT	35	2.7143	.8684
REWRITE	32	3.1408	.5834
ENGLORAL	39	4.7821	.3590
ENGLWRIT	40	4.5875	.6293
FRENCH	31	2.9355	.7471

Table 25

Corrrelations between Measures of Portuguese and English Written Sociolinguistics and Self-rated Proficiency, for the Sample Excluding Recent Arrivals

	PGLOBAV	PLETGLOB	PDIFFORM	PORTORAL	PORTWRIT	EGLOBAV
PLETGLOB	.9641**					
PDIFFORM	-0.0060	.0512				
PORTORAL	.1773	.2917	-0.0493			
PORTWRIT	.2469	.3439	-0.0541	.7515**		
EGLOBAV	.2428	.1720	.6137*	.1678	.3023	
ELETMIT	.3570	.3667	.5435*	.1291	.1943	.7440**
REWRITE	.3158	.2334	.7127**	-0.1437	-0.2205	.4347
ENGLORAL	.4008	.4298	.3372	.3448	.3858*	.2706
ENGLWRIT	.1185	.2326	.3882	.4548*	.3752*	.0522
FRENCH	.4588	.4457	.2816	.4108	.3339	.2707

	ELETMIT	REWRITE	ENGLORAL	ENGLWRIT
REWRITE	.3883			
ENGLORAL	.2689	.1703		
ENGLWRIT	.0810	-0.0992	.4590*	
FRENCH	.1610	.0184	.4227*	.5123*

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Table 26

Independent Variables Used in Regression Analyses of Portuguese Proficiency and Self-rating Variables for the Main Sample (Oral tests and Discourse Multiple Choice)[1]

[1] X in column 1 indicates a variable not used in the Grammar Multiple Choice regression analyses

X in column 2 indicates a variable not used in the Written Sociolinguistic regression analyses.

1	2	Variable name	Question no. in Use (U) or Attitude (A) Questionnaire	Description
		AOA	U:4	Age on arrival in Toronto
		FSJOB	U:12	Father's job: Pineo-Porter scale (1=High; 20=Low)
		PFAMUSE4	U:33a-f; 34a-f	Family use of Portuguese: =1 if there is a family member with whom student uses only Portuguese; =0 otherwise
		MEDIAP1	U:39,41,43	Use of Portuguese media: Books, Radio, Letters (1=Low; 5=High)
	X	TVHRSP	U:37	Hours/week watching Portuguese TV
X	X	VISITS	U:56	Number of visits to Portugal in last 5 years.
		MASS	U:49	Language used at Mass when student attends (1=Portuguese always, 5=English always)
		COMMEVNT	U:50	Language used by student to others at special community events (1=Portuguese always; 5=English always)
		CLASSA	U:45,47	Hrs/week Portuguese classes at a club + Yrs non-Heritage Portuguese classes

1	2	Variable name	Question no. in Use (U) or Attitude (A) Questionnaire	Description
		INSTRP2	A:1,2,4,6,8,10	Instrumental Portuguese: Student's perception of the value of Portuguese for success in education, employment, and the community
		VITALP2	A:24-30	Vitality of Portuguese: Student's perception of whether Portuguese is strong and valued by people. (Poles reversed on Q.25)
		VITALF1	A:26-30	Vitality of French: Student's perception of whether people valueFrench
		ACCPTP2	A:11,12,13	Student's acceptance of or liking for Portuguese
		ACCPTF2	A:11,12,13	Student's acceptance of or liking for French
		IMAKFRNP	A:5	Importance of Portuguese for making friends
X	X	ICOMFAMP	A:3	Importance of Portuguese for communicating with family
		ICOMFRNP	A:7	Importance of Portuguese for communicating with friends
		ICOMPEOP	A:7	Importance of Portuguese for feeling comfortable with people
		LIKETVP	A:14	I like to watch TV in Portuguese
		LIKEFRND	A:16	To talk with friends I like to use (1=Portuguese always; 5=English always)
		LIKEFAM	A:17	To talk to family I like to use (1=Portuguese always; 5=English always
		LMIXG	A:37,38	Acceptability of using Portuguese words when speaking English, and vice versa

1	2	Variable name	Question no. in Use (U) or Attitude (A) Questionnaire	Description
		GOODHOMP	A:32	"In my home people speak good Portuguese" (Disagree-Agree)
		EDMITJOB	A:41	My parents consider higher education more important than a job at age 16 (Disagree-Agree)
		HOMWORKP	A:45	My parents ensure I do Portuguese homework (Never-Always)
		KANDPP	A:50,52	Knowledge of and pride in Portuguese history and achievements
X		ORENTPX	A:47,48	Own and family's intention to stay in Canada (=1) or return to Portugal (=2)

Table 27

Independent Variables Used in Regression Analyses of Portuguese Proficiency and Self-rating Variables for the Main Sample (Oral tests and Discourse Multiple Choice)[1]

[1] X in column 1 indicates a variable not used in the Grammar Multiple Choice regression analyses

X in column 2 indicates a variable not used in the Written Sociolinguistic regression analyses.

1	2	Variable name	Question no. in Use (U) or Attitude (A) Questionnaire	Description
		AOA	U:4	Age on arrival in Toronto
		FSJOB	U:12	Father's job: Pineo-Porter scale (1=High; 20=Low)
		TALKWPAR	U:33c,d; 3434c,d	Language used in talking with parents (1=Portuguese always; 5=English always)
		MASS	U:49	Language used at MASS when student attends (1=Portuguese always; 5=English always)
		COMMEVNT	U:50	Language used by student to others at special community events (1=Portuguese always; 5=English always
		VITALEI	A:26-30	Vitality of English: Student's perception of whether people value English
		VITALFI	A:26-30	Vitality of French: Student's perception of whether people value French
		ACCPTE2	A:11,12,13	Student's acceptance of or liking for English
		ICOMFAME	A:3	Importance of English for communicating with family

1	2	Variable name	Question no. in Use (U) or Attitude (A) Questionnaire	Description
		LIKEFRND	A:16	To talk with friends I like to use (1=Portuguese always; 5=English always)
		LIKEFAM	A:17	To talk to family I like to use (1=Portuguese always; 5=English always)
		LMIXG	A:37,38	Acceptability of using Portuguese words when speaking English, and vice versa
		EDMITJOB	A:41	My parents consider higher education more important than a job at age 16
		KANDPE	A:51,53	Knowledge of and pride in Canadian history and achievements
		ORIENTPX	A:47,48	Own and family's intention to stay in Canada (=1) or return to Portugal (=2)

Table 28

Summary Multiple Regression Analyses for Self-Rated Proficiency Variables

Criterion	1st var.	2nd var.	3rd var.	4th var.	5th var.
PORTORAL	AOA	ACCPTP2	MEDIAP1	KANDPP	PFAMUSE4
df = 111	.00	.27	.30	.32	.34
PORTWRIT	AOA	MEDIAP1	KANDPP	ACCPTF	
df = 111	.00	.28	.37	.42	
ENGLORAL	AOA	ACCPTE2	TALKWSIB	ACCPTF	
df = 112	.07	.16	.18	.21	
ENGLWRIT	AOA	ACCPTF2			
df = 112	.03	.15			
FRENCH	AOA	ACCPTF2	FSJOB	ICOMFAME	LMIXG
df = 105	.06	.38	.42	.46	.47

Table 29

Summary Multiple Regression Analyses for Portuguese Proficiency Variables

Criterion	1st var.	2nd var.	3rd var.	4th var.	5th var.
POGITOTL	AOA	VISITS	CLASSA	VITALP2	COMMEVNT
df = 42	.02	.23	.30	.36	.41
PODGLOBL	AOA	VITALF1	CLASSA		
df = 42	.01	.15	.24		
PWDMULTC	AOA	CLASSA	TVHRSP	ORIENTPX	
df = 42	.00	.20	.31	.39	
POSDIFF4	AOA	FSJOB			
df = 42	.02	.11			

Table 30

Summary Multiple Regression Analyses for English Proficiency Variables

Criterion	1st var.	2nd var.	3rd var.	4th var.	5th var.
EPRPERAT df = 44	AOA .00				
ECUNITSZ df = 44	AOA .00	FSJOB .08			
EODDETL4 df = 44	AOA .00	TALKWSIB .10	ICOMFAME .16		
EWDMULTC df = 44	AOA .00	FSJOB .17	Mass .25		
EOSAVAPP df = 44	AOA .00	VITALE1 .05			

Table 31

Summary Multiple Regression Analyses for Portuguese and English Written Grammar and Sociolinguistic Proficiency Variables

Criterion	1st var.	2nd var.	3rd var.	4th var.	5th var.
PGRAMMC	AOA				
df = 24	.00				
EGRAMMC	AOA	LIKEFAM			
df = 28	.05	.16			
PDIFFORM	AOA				
df = 13	.00				
REWRITE	AOA	VITALF1	COMMEVNT	LIKEFRND	
df = 21	.05	.26	.43	.58	

Table 32

Comparison of Toronto and Azorean Performance on Portuguese Proficiency Measures

	Toronto		Azores		F value
	mean	SD	mean	SD	
PWDMULTC df = 1,128	.51	.16	.65	.11	32.8**
PODDETL3 df = 1,85	3.9	.74	4.4	.39	6.18*
PODGLOBL df = 1,85	3.0	1.20	4.0	.70	15.18**
POSDIFF4 df = 1,83	1.1	.30	1.2	.23	6.72*
POSAVAPP df = 1,86	4.5	.30	4.7	.19	17.02**
PDIFFORM df = 1,64	1.2	.38	1.2	.46	0.00
PGLOBAV df = 1,69	3.7	1.00	4.2	.54	7.25**
PLETGLOB df = 1,78	3.3	1.23	3.6	.91	2.52
PGRAMMC df = 1,107	.5	.18	.9	.06	326.20**

* p<.05
** p<.01

APPENDIX A

DESCRIPTION OF PORTUGUESE AND ENGLISH PROFICIENCY VARIABLES

1. SELF-RATINGS

PORTORAL	Portuguese Self-rated Oral Competence: Mean of "speak" and "understand"
PORTWRIT	Portuguese Self-rated Written Competence: Mean of "read" and "write"
ENGLORAL	English Self-rated Written Competence: Mean of "speak" and "understand"
ENGLWRIT	English Self-rated Written Competence: Mean of "read" and "write"
FRENCH	French Self-rated competence: Mean of "speak", "understand", "read", and "write"

2. ORAL TESTS, AND WRITTEN DISCOURSE

2:1 Portuguese variables

POGERRAT	(Errors in five categories) / (number of finite verbs)
POGEREXD	(Non-dialect errors) / (number of finite verbs)
PODCONCL	Oral Discourse: Total (max=2) for logic of 2 conclusions
PODDETL3	Oral Discourse: Mean of ratings of logic, anaphora, time
PODGLOBL	Oral Discourse: Impressionistic rating on 5-point scale
POGIGRAM	Oral Grammar: Proportion correct of 27 IDEA items
POGIPHON	Oral Phonology: Proportion correct of 9 IDEA items
POGITOTL	Oral "Grammar" Total: Proportion correct, 53 IDEA items
POGIVOCA	Oral Vocabulary: Proportion correct of 17 IDEA items
POSAPPHI	Oral Sociolinguistic "high": Mean appropriateness of markers of formality on 5 "formal" items
POSAPPLO	Oral Sociolinguistic "low": Mean appropriateness of markers of formality, 5 "informal" items

POSAVAPP	Oral Sociolinguistic: Average of appropriateness of markers of formality (low and high) items
POSDIFF4	Oral Sociolinguistic: Difference in four categories of markers of formality between low and high items
PPRPERRP	Prepositions: number of errors in picture description
PPRPTOKP	Prepositions: number (excluding 'a') used in picture description
PPRPTYPP	Prepositions: number of "types" (excluding 'a') in picture description
PPRPTYPS	Prepositions: number of "types" (excluding 'a') in story retelling
PVRBTNSS	Verb tenses: number used in story retelling
PWDMULTC	Discourse Multiple Choice: Proportion correct, 26 items

2.2 English variables

ECUNITSZ	C-unit: mean number of words per c-unit in story telling
EODANAPH	Oral Discourse: Rating of anaphora on 5-point scale
EODCNECT	Oral Discourse: Rating of connectors on 5-point scale
EODDETL4	Oral Discourse: Mean of ratings of logic, anaphora, connectors, setting the scene
EODGLOBL	Oral Discourse: Impressionistic rating on 5-point scale
EODLOGIC	Oral Discourse: Rating of logic on 5-point scale
EODSCENE	Oral Discourse: Rating of setting of scene on 5-point scale
EORLSYNT	Oral Syntax: Rating on 5-point scale, on story-telling
EOSAPPHI	Oral Sociolinguistic HIgh: Mean appropriateness of responses to "formal" variants
EOSAPPLO	Oral Sociolinguistic Low: Mean appropriateness of responses to "informal" variants
EOSAVAPP	Oral Sociolinguistic: Average of High and Low appropriateness
EOSDETHI	Oral Sociolinguistic High: Mean number (of 8 categories) of formal markers used in "formal" variants
EOSDETLO	Oral Sociolinguistic Low: Mean number (of 8 categories) of formal markers used in "informal" items

EOSDIFF	Oral Sociolinguistic: Mean difference between "formal" and "informal" in use of "formal" markers
EPRONUNC	Pronunciation: Rating on 5-point scale (story telling task)
EPRPERAT	Prepositions: Number of errors/obligatory contexts
EPRPTYTO	Prepositions: Number of "types"/number of correct tokens
EVRBERAT	Verbs: (Number of errors) / (number of verbs)
EVRBTYTO	Verbs: (Number of "types") / (number of "tokens")
EWDMULTC	Discourse Multiple Choice: Proportion correct, 32 items

3. WRITTEN GRAMMAR

PGRAMMC	Portuguese Grammar Multiple Choice: Proportion correct of 49 items
EGRAMMC	English Grammar Multiple Choice: Proportion correct of 34 items

4. WRITTEN SOCIOLINGUISTICS

3.1 Portuguese variables

PDIFFORM	Difference in use of formal markers: Letter - Note
PGLOBAV	Mean of appropriateness ratings of Letter and Note
PLETGLOB	Letter: Rating of Global appropriateness, 5-point scale

3.2 English variables

EGLOBAV	Mean of appropriateness ratings of Letter and Note
ELETMIT	Letter: Rating of appropriateness of use of mitigation
REWRITE	Mean rating of appropriateness of rewriting of phrase in short passages

APPENDIX B
SCORING SCHEMES

1. ORAL PORTUGUESE GRAMMAR

1:1 Story Retelling Task

The story retelling task is scored for the number of occurrences of each of the following:

Number of finite verbs. A finite verb includes any conjugated verb form, and excludes past participles, gerunds, and infinitives. The count is of finite verbs actually used (occurrences) rather than of obligatory contexts.

Number of errors. Errors in the following categories are counted:

(1) Errors in syntax. Syntactic errors include addition, omission, substitution, or misplacement within a sentence of a word or phrase. The types of word or phrase considered include, for example, the pronominal subject of a verb, reflexive particles, adjectives, and adverbs. They do not include verbs or prepositions, which are counted separately.

(2) Errors in verbs. Verb errors include use of the wrong tense, and errors in agreement of number or gender. Also included is the use of TER (a form which is not used in mainland Portuguese but current in Brazil and in some islands of the Azores) instead of HAVER. If á single verb form has errors in both tense and agreement, both are counted.

(3) Preposition errors. 'Preposition' includes both simple prepositions and prepositional locutions; it does not include a preposition used as part of an adverbial locution. An error is counted in case of the addition, omission, or substitution of a preposition, or the use of a word or phrase with a different grammatical function in place of a preposition.

(4) Errors in degree of adjectives. Any case of the incorrect use of the positive, comparative, or superlative form of an adjective is counted.

(5) Transfer errors. The use of any English grammatical construction, idiom, or colloquial expression is counted as an error. A single noun, verb, preposition (or prepositional locution), article, or conjunction used in English or 'adapted' from English is not counted.

(6) 'Dialect' forms. All 'errors' counted in the above categories which occur commonly in Brazilian or Azorean Portuguese are counted.

Measures of sophistication.

(1) Range of prepositions, excluding 'a'. Correctly used, different prepositions and prepositional locutions ('types') are listed and counted. ('A' was not counted because of its extremely high frequency in normal use of Portuguese.)

(2) Range of verb tenses. The different verb tenses correctly used are listed and counted.

1:2 Picture Description Task

The picture description is scored for prepositions only, in five categories. 'Preposition' includes both simple prepositions and prepositional locutions; it does not include a preposition used as part of an adverbial locution.

(1) Number of occurrences. Each use of a preposition is counted. The count does not include omitted prepositions, or any other grammatical form substituted for a preposition.

(2) Occurrences, excluding 'a'. This is the same as for the preceding count, except that any occurrences of the preposition 'a' are ignored.

(3) Errors. An error is counted in case of the addition, omission, or substitution of a preposition, or the use of a word or phrase with a different grammatical function in place of a preposition.

(4) Range. The number of different, correctly used prepositions (types) is counted.

(5) Range, excluding 'a'. This is the same count as for Range, except that the preposition 'a' is disregarded if it occurs.

2. ORAL PORTUGUESE DISCOURSE

The story retelling is scored on a five-point scale for each of the following three aspects of discourse competence.

Logical sequence. A rating from 1 to 5 is assigned according to the following criteria:

- 5 Story retelling is logical and coherent.
- 4 Some fragmentation, change of topic, or disordering of events, but the narrative is still coherent.
- 3 Some fragmentation, etc; the narrative is slightly incoherent.
- 2 Some fragmentation and omission of important information.
- 1 Sentences do not connect; no coherence.

Anaphora. A score from 1 to 5 is assigned on the following criteria:

- 5 Subsequent references (especially antonomiasis and direct and indirect object pronouns) to the cat and Louise are appropriately made; gender is correct.
- 4 Certain pronouns are omitted but gender of those used is correct.

- 3 The gender of one character (but not both) changes throughout the story, but subsequent references to characters are given.

- 2 Same as (3) but some of the subsequent references to characters are omitted or imprecise.

- 1 No appropriate references; the gender of both characters changes throughout the story.

Time orientation. The following guidelines are used to assign a score from one to five.

- 5 Sequence of tenses across sentences defines past and present tenses appropriately; or, the story is retold in the present tense by appropriate use of temporal adverbs or conjunctions.

- 4 Same as (5), but student shows some problem in the use of tenses or temporal adverbs or conjunctions.

- 3 Locating of events in the past or present is appropriately conveyed through the use of tenses, but aspect or mode is not correctly used (e.g., imperfect instead of perfect, or indicative instead of subjunctive) across sentences.

- 2 Past and present tenses do not refer appropriately to past and present events in the story.

- 1 Tenses are so incorrectly used as to make the sequence of events totally illogical; notions of aspect and tense are not present.

3. ORAL PORTUGUESE SOCIOLINGUISTIC TEST

3:1 Detailed Scoring

Each of the twelve situation-variants (items) in the Portuguese oral sociolinguistic test is scored for level of formality in the use of five categories of marker, as described below. In addition, the entire item is coded as missing if the student fails to respond (code=9), gives an off-topic or otherwise uncodable response (code=8), or says that he or she would say nothing (code=6).

Initial form of address. A three-point scale is used:

Score or code	Level of formality	Examples
3	High	'O senhor' or 'A senhora' - with or without name or profession
2	Medium	Você
1	Low	'Tu' or 'Eh pá' or the absence of either when the verb is in the second person singular
code 7	'missing'	Use of no form of address

Use of person in subsequent pronouns and verb agreement. Under this heading, any initial form of address is disregarded. In Portuguese, the third person singular is a formal mode of address compared to the second person. Scoring therefore focusses on reflexive, possessive, direct, and indirect object pronouns; and verb agreement - i.e., use of third vs second person. Use of verbal agreement involving other than second or third person is considered as a grammatical error and disregarded.

Score or code	Level of formality	Examples
3	High	Use of the third person singular, only; e.g., 'A senhora podia ajudar-me no Português?'
2	Mixed	Use of both third and second person; e.g., 'Por favor senhor bibliotecário, pode-me (third person) emprestar a tua (second person) régua?'
1	Low	Use of second person singular, only; e.g., 'Miguel, podes-me emprestar a tua régua se faz favor?'
code 7	missing	neither pronouns nor verb agreement is used.

Use of modality. This category concerns the use of modality in verbs such as PODER (can, to be able to), QUERER (to want, to wish) DEVER (should) and IMPORTAR-SE (to mind) functioning as markers of higher formality. However, in Portuguese, verbal aspect (e.g., imperfect or conditional vs indicative) and modality (e.g., subjunctive vs imperative) are important means of attenuating the abruptness of directives; therefore, scoring is not restricted to the above verbs.

Score or code	Level of formality	Description and Examples
3	High	Use of modal verb with or without aspectual attenuation, e.g., 'Quer que lhe traga o programa do espectáculo?' 'Podes-me deixar usar a régua?'
2	Medium	Use of non-modal verbs such as DAR (to give) in the imperfect or conditional but without one of the above modal verbs. e.g., 'A senhora por favor falava mais baixinho que estou a fazer um teste?'
1	Low	Neither modal verb nor modal or

aspectual attenuation is used.
e.g., 'Paula, fala mais baixo!'
'Oh Paula empresta-me aí a régua?'

Extra politeness markers. This category deals with fixed expressions of politeness; use of one or more such expressions, or expressions that in a given context function as extra politeness markers, is scored as 3, while absence of one is scored as 1.

Score or code	Level of formality	Description and Examples
3	High	One or more expression such as - 'Por favor' (please) - 'Se faz favor' (if you please) - 'Com licença' (excuse me, pardon me) - 'Desculpe/desculpa' in initial position (sorry, excuse me) - ' senhora quer vir para aqui? Pode-se ver melhor.'
1		absence of any extra politeness marker

3:2 Rating of Appropriateness

Each item is given a rating from 1 (completely inappropriate) to 5 (completely appropriate) for its overall appropriateness as a response to the situation and addressee presented in the item. This is done by the rater on a subjective basis, keeping in mind that a high score on formality for use of markers would correspond roughly to a high score for appropriatenes on 'formal' variants, but to a low (inappropriate) score on 'informal' variants.

4. WRITTEN PORTUGUESE SOCIOLINGUISTIC TEST

4:1 Detailed Scoring

The letter and note in Portuguese are each scored for formality in each of the following categories. In general, sentences in English are ignored; this may result in a missing data code for the category (see, for example, Opening). A letter or note which is completely off topic, or which is too short to provide sufficient data for scoring, is coded as missing for all categories.

Opening. The opening potentially consists of two parts, a form of address and a title, each of which may occur in a formal or an informal variant. Examples are:

Formal/ Informal	Address	Title
Formal	'Caro(a)' (dear)	'Senhor(a)' (sir/madam)

Formal	'Exmo.(a.)' (dear')	'Departamento de Turismo'
Informal	'Querido(a)' (dearest)	'Amigo(a)' (friend)

The opening is then given a score from 1 to 3 as follows:

Score or code	Level of formality	Example
3	High	Formal address and formal title
2	Medium/mixed	Either a formal address or a formal title but not both. The other may be informal or not used.
1	Low	Both address and title are informal, or a formal title but not both. The other may be informal or not used.
code 7	missing	Neither address nor title is used; or, the opening is written in English.

Use of person in subsequent pronouns and verb agreement. Under this heading, any initial form of address is disregarded. Scoring focusses on reflexive, possessive, direct, and indirect object pronouns; and verb agreement - i.e., use of third vs second person. Use of verbal agreement involving other than second or third person is considered as a grammatical error and disregarded.

Score or code	Level of formality	Examples
3	High	Use of third person singular, only; e.g., 'Os senhores podiam enviar-me catálogos sobre o Algarve?'
2	Mixed	Use of mixed third and second person; e.g.,'Por favor, senhor, podes-me (second person) enviar informação sobre o seu (third person) programa de Verão?'
1	Low	Use of second person singular, only; e.g., 'João. Podes arrumar o teu quarto?'
code 7	missing	neither pronouns nor verb agreement is used; the student fails to address the person to whom he or she is writing.

Use of modality. This category concerns the use of modal verbs such as PODER (can, to be able to), QUERER (to want, to wish), DEVER (should) and IMPORTAR-SE (to mind) functioning as markers of higher formality. Verbal aspect (e.g., imperfect or conditional vs indicative) and modality (e.g., subjunctive vs. imperative) are other means of attenuating the abruptness of directives; therefore, scoring also takes into account the use of such attenuation in non-modal verbs. Modal verbs, however, are considered to contribute more to formality, and this is taken into account in scoring. The text of the letter or note is assigned a score from 1 to 3, based on the number of modal and attenuated non-modals used, as shown in the following table:

Score	Number of modals	Number of attenuated non-modals
1	0	0
1	0	1
2	0	2 or more
2	1	0
3	1	1 or more
3	2	0 or more

Extra politeness markers. This category deals with fixed expressions of politeness; use of one or more such expressions, or expressions that in a given context function as extra politeness markers, is scored as 3, while absence of one is scored as 1.

Score or code	Level of formality	Description and Examples
3	High	One or more expression such as - 'Por favor' (please) - 'Se faz favor' (if you please) - 'Com licença' (excuse me, pardon me) - 'Desculpe/desculpa' in initial position (sorry, excuse me) - 'Por favor, podia-me enviar catálogos...'
1		absence of any extra politeness marker

Closing. The closing of the letter or note is rated from 1 to 3, using the following examples as guidelines.

Score or code	Level of formality	Description and Examples
3	High	'Aguardando resposta' (waiting for your reply) 'Atenciosamente' (Truly, Sincerely)

		'Muito obrigado(a)' (Thank you very much)
2	Medium	'Obrigado' (Thanks, Thank you)
1	Low	'Beijos' (kisses) 'A mãe' / 'A mamã' (Mother/mommy) 'Adeus' (Bye)
		Signature only
code 7	missing	Neither closing nor signature used.

4.2 Rating of Appropriateness

Each letter and note is given a rating from 1 (completely inappropriate) to 5 (completely appropriate) for its overall appropriateness given the person to whom it is addressed and the content required. This is done by the rater on a subjective basis. In general, the rating should be expected to reflect the appropriateness of use of more markers for formality in the letter and fewer in the note.

5. ORAL ENGLISH GRAMMAR

5:1 Communication Units

Scoring is done from transcripts. A Communication Unit (C-unit) is defined as a clause and its modifiers (Gambell 1978). The boundaries of each C-unit are marked, and the C-units are counted. The number of words in each C-unit is counted, and a total obtained. The count of words excludes all extraneous material (e.g., transition markers such as 'umm'; repeated words or phrases; the first attempt in a self-correction; false starts; and words which perform no function in the C-unit per se, such as 'okay?', 'you know?'). Contractions such as 'it's' count as two words.

In the following examples, extraneous material is enclosed in brackets; the number of words in each C-unit is shown at the beginning of the C-unit.

4 (So one -) it was night time, (okay?)

6 And he's (he's) always fooling around

9 And now (he has no) the neighbours don't talk to him

8 He was trying to explain (um) at the cat

15 (Well) (um) the neighbour's neighbour which is (the cat's) the house that the cat's on (um) (he) came out

5.2 Verbs

Only finite verbs are scored; infinitives and participles are not. For example, in 'he tries to get', 'tries' is counted; 'to get' is not. The verb 'to be' is not counted, because of its extremely high frequency in English. The number of finite verbs used (tokens) and the number of obligatory contexts for use of a finite verb are counted. Errors and obligatory contexts are scored on questions 7, 8, and 9. Sophistication (types and tokens) are measured on question 14 (story telling).

Errors. Only gramm[illegible]al, not lexical, errors are scored. The following are counted as errors:

- use of an incorrect tense
- omission of the verb
- omission of the auxiliary
- omitted third person singular /s/

Range of verbs ('types'). The number of different verbs occurring in a finite form is counted.

- Copula is not counted.
- A verb used in different senses is counted only once: in, 'he gets some water', and 'he gets mad', 'gets' is credited as one type only.
- A verb occurring with a prepositional adverb in a verb phrase is in some cases counted as a separate type: 'he goes up to the roof', 'he goes in' are counted as the type 'to go'; while 'he goes back' and 'he goes away' are counted as the distinct types, to go back and to go away.
- Lexical correctness is not considered: in 'he puts the water at the cat', 'puts' is counted as a type.

5.3 Prepositions

Prepositions are scored from the transcript of the story telling task.

Definition. The definition of a 'preposition' is based on Quirk et al (1972) and includes both simple and complex prepositions. The main forms of complex preposition are preposition-preposition, preposition-noun-preposition, and adverb-preposition. To be considered a complex preposition, lexical or syntactic coherence is required; such coherence is understood to demand some subjective judgement by the scorer.

The following are excluded from the preposition score:

- to in the infinitive form.
- a preposition in a fixed adverbial locution; e.g., once upon a time, by mistake, first of all, after all, at worst, before long.
- a preposition in a verb phrase in which the verb-preposition sequence forms a lexical unit essentially different in meaning from the verb alone; e.g., to go on, meaning 'to continue'; to get off on, meaning 'to enjoy'.

- a prepositional form used adverbially; e.g., the car drove past; the house is beside.

NOTE, however, that a stranded preposition is counted as correct; e.g., on, in 'the house which the cat is on'.

In general, the student's basic construction is accepted whenever possible, and usage is judged within the context produced. For example, 'the house where the cat is on' could be understood as 'the house where the cat is' or 'the house which the cat is on'. In this case, on is accepted as a correct stranded preposition.

Prepositions are scored for errors, and four scores are given to each student:

1. Total number of errors
2. Number of obligatory contexts for use of a preposition
3. Number of correctly used different prepositions (types)
4. Number of prepositions correctly used (correct tokens)

Errors. Some general guidelines are followed. In particular, a preposition, whether simple or complex, is never counted as having more than one error; e.g., 'at top on' instead of on top of is one error. Errors are judged fairly strictly against standard English usage, following Quirk et al. However, leniency may be required when the error may lie not in the choice of preposition but in some other part of the context, such as the choice of a verb. For instance, 'He put the water at the cat' could mean 'He threw the water at the cat' or 'He put the water on the cat'. Given the set of pictures on which the story is based, it is usually clear that the error is in the verb rather than the preposition, and no error is counted. Finally, in the case of self-correction of a preposition or an entire phrase, the final version is scored; in the case of exact repetition, only one is scored.

The following are the types of errors counted:

(1) omission: failure to use a preposition where one is required; e.g., 'the lady ( ) the first house'.

(2) addition: use of an additional preposition where it is (a) syntactically incorrect; e.g., 'the room at downstairs'; 'he came to home'; or (b) lexically incorrect; e.g., 'the man from downstairs' (meaning, 'the man who is downstairs').

(3) substitution: use of a lexically incorrect preposition; e.g., 'He spilled the water to the man'; 'The cat was on top of (on) the roof; 'The man was under the house'.

5:4 Oral Syntax and Pronunciation

These measures are scored from the tape for the story telling task, immediately after it has been scored for oral discourse. The tape is not replayed unless necessary. A subjective judgement is made, first of oral syntax, then of pronunciation, using the following five-point scale.

(1) intelligible only in set expressions

(2) intelligible at times

(3) completely intelligible

(4) almost native-like

(5) native-like

6. ORAL ENGLISH DISCOURSE

Oral competence in English discourse is scored from the tape of the story-telling task. Grammatical errors are ignored in this scoring. The tape is played, and a score is assigned on global discourse competence. Detailed scoring is then conducted, with the tape being replayed only if necessary.

Global rating. The rater gives an impressionistic rating of the overall performance of the student on discourse, using the following scale:

(1) completely disjointed

(2) somewhat disjointed

(3) basically connected

(4) well connected

(5) very well connected

Setting the scene and indentification of characters, etc. Relevant features of the story involve:

- it is night and people are sleeping
- the cat is making noise
- there is a roof

Of primary importance are windows, jug, water, upstairs man, lady, and downstairs man. Of secondary importance are moon, curtains, pyjamas, and apartments. Use of the definite and indefinite article is considered in scoring. A score is assigned on the following scale:

(1) setting/characters/objects unidentified (articles omitted)
(2) setting/characters/objects identified poorly (articles omitted)
(3) setting/characters/objects identified only partially (articles unclear)
(4) setting/characters/objects almost completely identified (articles cohesive)
(5) setting/characters/objects completely identified (articles cohesive)

Logical sequence of events. A series of events should be reported and their temporal relationships made clear. In the following list, essential events are marked with an asterisk.

- * The cat is sitting on the roof at night.
- * The cat is meowing, singing, making noise.
- The upstairs man opens the window and screams (yells) at the cat.
- The cat stops making noise.
- The upstairs man goes back in.
- The cat starts to make noise again.
- The upstairs man comes out again with a jug of water.
- At the same time, a man downstairs in another window puts his head out.
- * The upstairs man throws water at the cat, but he misses and hits the downstairs man.
- The two men yell at each other, and
- * A woman downstairs puts her head out (to see what's happening).
- * The upstairs man accidentally bumps the pitcher (jug) and spills the rest of the water on the woman.
- The upstairs man is unhappy (says something), but
- * The cat keeps on making the noise.

General considerations are whether the story is consistently sequenced in the present simple and present continuous or in the past simple and past continuous; whether it gives incorrect information (such as, 'the man drops the vase on his wife'); and gives acceptable implications or completion for the story.

A rating is assigned on the following scale:

(1) no sequencing of main or secondary events

(2) only some main events sequenced; false information given

(3) only some main events sequenced

(4) all main events sequenced clearly

(5) all main events and all secondary events sequenced clearly

Anaphora. This is concerned with correctness and clarity in the use of pronouns (he, she, it), and adjectives (that man, the other guy), to refer to characters and objects. Incorrect, ambiguous, or confused references result in a lower score on the following five-point scale:

(1) reference incorrect to antecedents

(2) reference confused to antecedents

(3) reference marginally correct and clear to antecedents

(4) reference almost correct and clear to antecedents

(5) reference correct and clear to antecedents

Logical sentence connectors. In this category, cohesive devices which establish the sequence of and relations among events are considered, especially such connectors as 'first', 'then', 'so', 'and so', 'because', 'instead of'. Use of only 'and... and...' or 'the first picture... the second picture...' is considered as listing events. A score is assigned on the following scale:

(1) events listed or connected with one type of time connector

(2) events connected with few time connectors

(3) events connected with variety of time connectors

(4) events connected with time and causal connectors

(5) events connected with time, causal, and contrastive connectors

7. ORAL ENGLISH SOCIOLINGUISTIC TEST

Scoring is done from transcripts; all students are scored on each item before the next item is begun. Scoring of global appropriateness is done after completion of all detailed scoring.

7:2 Description of Transcripts

Each answer is transcribed in a simplified version:

- Where relevant to deciding which part of a text should be scored, if any, comments by or discussion between the interviewer and a student are included, but enclosed in ( ).
- Extra-linguistic material (hesitations, self-corrected phrases, pause-fillers, etc) is omitted.
- Contractions are rewritten in full standard English, except when they involve more or less standard forms such as auxiliary verbs, negations, etc. Thus,

'wanna' is transcribed as want to, 'cuz' as because, but 'hasn't', 'I'll' are written as spoken.

- Contractions involving a feature of the scoring scheme, especially 'can', 'could', 'will' and 'would', are transcribed as accurately as possible, and followed by a semi-phonetic transcription where some doubt remains. (NOTE: In scoring, doubtful features are not given points.)

- Rising, falling, and level intonations are marked at the end of each answer and at other possibly relevant points in the text. Since students varied in their ability to role-play in the text, these indications are intended mainly to clarify ambiguous forms such as uninverted questions which resemble imperatives.

7:3 General Notes

(1) A complete or partial repetition was occasionally asked for when the interviewer thought that the recording might be inaudible. If the original is understandable it is scored, because any changes in the second version may be a result of the student's thinking the request for repetition meant that the first was incorrect. If it is incomplete or garbled, the second may be taken into account.

(2) A marker for which the transcript shows an unclear reading is not given a point.

(3) Missing data codes are assigned when (a) there is no response; (b) an item is unscorable: incomplete; off-topic; unclear; etc.; (c) no explanation is provided (see p. xxx, below); and (d) a student says that they would say nothing in the situation.

(4) A phrase may be given more than one point if it incorporates formal markers from different categories. For example: 'if you wouldn't mind' is given 3 points: for conditional clause, modal auxiliary, and politeness marker (see below).

7:4 Detailed Scoring of Linguistic Features

The response to each of the twelve situational variants (6 formal and 6 informal) is checked for use of formal features in eight categories. For each category in which one or more formal features occur, a single point is given. The maximum formality score is thus 8. Non-use of a category, or use of only informal, neutral or non-native-like features in that category, is scored as 0, except in Explanation in which a 'missing data' code is assigned. Under each category below, examples of formal features appear in the left, while features for which no point is given are on the right.

Interruptor. A word or phrase whose main function is to catch the attention of the addressee, often involving an apology for interrupting whatever the addressee was doing:

pardon me	hey
excuse me	say
sorry for bothering you	hello, hi

good morning

Address. A title or name, or other word or phrase used to name the addressee:

sir	mister
ma'am, madam	missus, lady
officer	librarian
	personal names (Ann, Bob, Mr. Warren)

Politeness marker. Set expressions such as:

please	it's yours at will
thank you, thanks	absence of marker
expressions with 'mind' (e.g., Do you mind; If you wouldn't mind)	

Modal auxiliary. This is scored when involved in the directive itself, rather than, say, in an Explanation. In 'You can sit here so you could see better', can is part of the offer, while could is part of the Explanation; the score for the modal auxiliary can is 0.

could	can, will
would	may you, you may
may I (asking permission)	

Grammatical mood. Mood of the verb(s) in the speech act, as indicated by the structure of the sentence, the intonation, or the meaning:

interrogative (inverted)	partial/intonational questions (e.g., You mind? You want me to do it?) declarative imperative

Conditional clause. A clause which reduces the force of the speech act by suggesting that the addressee has some freedom to refuse to comply, because of some condition, including his or her preference:

if you have some time	do you have some time
if you want	do you want
if you wouldn't mind	
if you have room	do you have room
if you don't need it	do you need it
would you like it if I did it	
I was wondering if	

Attenuation. A point is given for use, or zero for absence, of a word or phrase which reduces the force or restricts the scope of a directive or offer by:

(1) reducing the amount or kind of effort or inconvenience involved:
a little bit, for a while
some, sometime, somewhat,

when it's convenient
I'll give it back; just to...
give me the notes (rather than 'help me')

(2) providing information, such as a specific time or place, on the basis of which the addressee can more easily make a decision:

can you help me today?
can I stay in at recess?

(3) reducing the debt incurred by the addressee if s/he were to accept the offer:

gladly
it's my pleasure
it's no trouble
no-one's using it
you can pay (for the pen)

(4) reducing the responsibility of the addressee for an (offending) action:

can you possibly
use of an apology rather than a complaint

Explanation. Give some explicit reason (usually drawn from the stimulus tape) for a directive, which gives the addressee grounds for evaluating the validity of the request, complaint, etc., and therefore deciding whether to comply. Failure to use a marker in this category is scored as missing data rather than 0, because students were directed not to give too much detail, and because much of what is scored in this category is provided in the stimulus tape and may therefore be omitted for reasons (such as inattentiveness or forgetting) unrelated to sociolinguistic competence.

	missing data
there's a test coming I've been away I'm really having trouble to see better no-one is using it I overheard (that you need a pen) it's close to my house I'm trying to study you may get thrown out someone may get hurt	absence of a marker (Note: I need it or You're bugging me are not considered to be explanation.)

Impressionistic Rating of Linguistic Appropriateness

Each answer is scored on a 5-point scale for appropriateness of language use. The rater reads the text and assigns a score according to his or her impression of how appropriate it is, by native-speaker standards, given the speaker, addressee, and other circumstances described in the test. The scale is treated as an absolute judgement rather than as a ranking of the students' answers: the best in the set for a given item may or may not merit a 5, and the worst may not be completely inappropriate; thus it is possible for all answers to a particular item to be assigned the same score.

The scale is from
1 = completely inappropriate
to
5 = completely appropriate

8. WRITTEN ENGLISH SOCIOLINGUISTIC TEST

There were three tasks presented in the Written English Sociolinguistic Test, requiring the writing of a letter requesting Environment Canada for information; the writing of a note from a teacher to a negligent student about homework; and the rewriting of a phrase in each of six paragraphs.

Detailed scoring of the letter and note did not in general attain an adequate level of reliability because of small amounts of data and small numbers of cases. The following scheme includes criteria only for those measures discussed in this report.

Letter: Appropriateness of mitigation. A native speaker may adjust the appropriateness of the phrasing of a directive to different addressees by mitigating its directness in several ways, such as:

- use of modal verbs; e.g., could, would, may, might;
- use of conditional clauses; e.g., 'if you wouldn't mind', 'if at all possible';
- addition of politeness markers; e.g., please, thank you, or set phrases such as 'would you be so kind';
- use of a question form.

The scorer makes a subjective evaluation of the appropriateness of use of mitigation in the body of the letter, based on the use of these and other techniques, and assigns a score on a five-point scale from 1 = completely inappropriate to 5 = completely appropriate.

Global appropriateness. The letter and the note are each assigned a rating from 1 (inappropriate) to 5 (completely appropriate) for global appropriateness, on a subjective basis within the following constraints:

- Letter: The opening and closing are disregarded. The body of the letter is read and judged impressionistically on how closely it approximates native-writer standards of appropriateness, overall The rating may be open to influence by any linguistic feature which affects appropriateness.

  Note: Any opening or closing is disregarded. The body of the note is judged impressionistically on how closely it approximates what a teacher might write. This score may be open to influence by any linguistic feature which affects the appropriateness, except the extent to which it incorporates or adds to the information provided in the instructions. Emphasis is not on how much detail is written, but on how it is presented.

Rewriting phrases in short paragraphs. The test comprises six short passages, each with an underlined phrase in which the style differs from that of the rest of the text. The task was to rewrite this section so that the style 'fit better' with the rest, while maintaining the meaning of the original phrase.

Scoring focusses on the improvement in the sociolinguistic consistency of the passage, in terms of the following:

- register; e.g., formal, casual, business-like, professional;
- writer (apparent or explicit); e.g., mother, teacher, employer, journalist;
- audience (apparent or explicit role of the recipient); e.g., child, student, employee, reader of magazine;
- format of the communication; e.g., personal letter, business letter, magazine article, memo;
- purpose of the communication;
- content.

Grammatical errors and minor deviations from the original meaning are disregarded, unless they directly affect the appropriateness.

A rewrite is considered unscorable and coded as missing data: (a) if handwriting or spelling make the response unreadable; (b) if the response is off-topic or otherwise shows that the student failed to understand the nature of the task; (c) if the response is not comparable to the original phrase because of incorrect vocabulary or syntax, or because the meaning is otherwise too different from the original; or (d) if the student fails to attempt the item.

An item considered scorable is then judged impressionistically on how much it improves on the original phrase in approximating a native-like appropriateness. Two short items are rated on a 3-point scale; the other four on a 5-point scale:

Items 1,5:

Rating	
1	No improvement
2	Moderately appropriate
3	Completely appropriate

Items 2, 3, 4, 6

Rating	
1	No improvement
2	Somewhat appropriate
3	Moderately appropriate
4	Almost appropriate
5	Completely appropriate

APPENDIX C

1. RELIABILITY OF PORTUGUESE PROFICIENCY MEASURES

1:1 Oral Portuguese Grammar

IDEA Test. Scorer L explained the scoring to scorer R, who then re-scored 21 students as a check on reliability. Of the 1155 (55 x 21) items re-scored, there was agreement on 1089, or 94.3%. The scorers reported that two items were often unscorable: one elicited mixed English and Portuguese forms; the other could be answered without a verb (for which the item was to be scored). These two, which accounted for 11 of the 66 disagreements, were excluded from the analyses. On the remaining 53 items, there was 95.1% agreement.

Expressive language measures. Following scoring of the story retelling, three students were selected at random from each of the seven classes, and were rescored by the scorer who had not previous done them as a check on reliability. T-tests of the difference between the means of the two scorers and the correlations between their scores were computed.

On the picture description task, scorer R scored all seven schools for number of prepositions occurring (excluding 'a'), range of correctly used prepositions (excluding 'a'), and errors in prepositions. Scorer L then rescored 21 students (three from each class) and reliability was investigated.

Inter-rater reliability for expressive language measures of oral Portuguese grammar is reported in Table C.1. Measures for which no reliability is reported had near-zero variance, or were scored by the two scorers working together (range of tenses).

1:2 Oral Portuguese Discourse

Story retelling. Scoring procedures were developed and revised for the three detail measures of oral Portuguese discourse. Final inter-rater reliability for 21 students scored by both scorers is reported in Table C.2.

Global scoring. Following the scoring of five classes, two classes - one scored originally by each scorer - were rescored by the other scorer. Twenty-three students were thus used to check the inter-rater reliability. The level of agreement was high (see Table C.2). The remaining two classes were then scored; no further reliability check was conducted.

Logical conclusions. Three students from each class were rescored to provide an estimate of reliability. Scoring allowed for three codes for incorrect answers: repetition of an earlier part of the story, other inappropriate responses, and non-response, and for a single point for each response which made sense as a conclusion to the story. The two scorers agreed on 85% of the 42 (2 x 21) cases scored or coded. The total score for the two items showed no significant difference between the scorers (mean difference = .05; t = .57) but a correlation of .91 (p .01).

1:3 Oral Portuguese Sociolinguistic Test

Development and training with the scoring scheme involved the two scorers jointly scoring one class. One scorer then scored three classes, the other four. They then each rescored three students from each of the other's classes, to check their reliability. Computation was done across all 252 items (21 students x 12 situation variants). Tone of voice involved similar training and checking. Finally, the response to each item was given a global rating of appropriateness, following similar procedures.

Inter-rater reliability for the four sociolinguistic details and global rating is shown in Table C.3. Numbers of cases differ because items coded as unscorable for any reason could not be included. The reliability check on sociolinguistic details revealed disagreements on what constituted an 'unscorable' item. Each scorer then rechecked both his own and the other scorer's entire set of data for such discrepancies. On the final set of 4080 datapoints (12 items x 85 cases x 4 variables) there was agreement on all but 34.

1:4 Written Portuguese Discourse (Multiple Choice)

Inter-item reliability of the written Portuguese discourse test was carried out, for all students ($n = 85$, $\alpha = .700$) who completed the test.

1:5 Written Portuguese Grammar (Multiple Choice)

In the 49 items of the test, six subtests were identified: agreement of adjectives (4 items); articles (4); prepositions (13); agreement of verbs (7); tense of verbs (9) and word order (12 items). Inter-item analysis of the reliability of the written Portuguese grammar test was conducted on the total test, and on the subtests. For the 47 students who completed the test, results are reported in Table C.4.

1:6 Written Portuguese Sociolinguistic Test

After all letters and notes had been scored for use of five sociolinguistic features, three students were selected from each class and rescored by the other scorer. Six letters and ten notes were coded as non-scorable (missing or off-topic) in total by both scorers. For the remaining letters ($n = 15$) and notes ($n = 12$), agreement was good for most of the scores, although scoring of use of modality showed enough disagreement that it was subsequently clarified and rescored for the letters. Following detailed scoring, global appropriateness was scored, and the reliability checked in the same way.

T-tests and correlations were computed across all scorable letters and notes (rather than separately for letters and notes), but should be interpreted cautiously because of the small number of cases and the small range of values (1 or 3 for politeness markers; 1 - 3 for the other four detail scores; 1 - 5 for global appropriateness). Results are in Table C.5.

1:7 Oral English Grammar

For certain of the oral English grammar measures two scorers either worked together in doing the counting, or checked each other's counts; for these, no reliability is reported.

Prepositions. Fourteen students were rescored by a second native speaker for their preposition errors and sophistication. Specific counts were made of (1) the number of errors made, (2) the number of obligatory contexts for prepositions, (3) the number of different, correctly used prepositions (types), and (4) the number of correctly used prepositions occurring (correct tokens). Inter-scorer reliability for each, as well as for the ratios errors/obligatory contexts and types/correct tokens, is shown in Table C.6. In addition, it was found that of 159 prepositions (or prepositional phrases) listed by one or both scorers, there was complete agreement on both the form of the preposition and its scoring in 113 cases (71%).

Global syntax and pronunciation ratings. For these two measures, the raters did the assessments independently, but compared scores on every five to ten students to ensure adherence to the scoring criteria. The mean of their ratings was used in analysis. Final inter-rater reliability for each measure is reported in Table C.7.

1:8 Oral English Discourse

The two raters, working independently but with regular comparison of their assessments, made impressionistic judgements of global discourse and four details. Disagreements of two or more points were discussed to ensure adherence to the scoring criteria. Inter-rater reliability for each of the variables is presented in Table C.8.

1:9 Oral English Sociolinguistic Test

An initial list of categories of markers was developed, including:

- Form of address
- Politeness marker
- Modal auxiliary
- Conditional clause
- Mood (inversion)
- Explanatory information
- Attenuation

Inter-rater reliability on these categories was conducted with a sample of 16 transcripts. Two native speakers scored each of the twelve items (situational variants) for all 16 students on the above categories. For each item, one point was given in each category for which at least one formal marker was used. Table C.9 suggests a high degree of agreement between scorers in each category. In the course of scoring, it was found that certain phrases such as excuse me which had been included under attenuation fit more appropriately in a new category of interruptor. The categories were revised to allow for this change. Final detailed scoring was carried out by one native speaker, using the revised categories; because of the high level of agreement in the preliminary study, no further inter-rater reliability check was carried out.

Subsequently, the items were rated for appropriateness on a five-point scale. Inter-rater reliability for appropriateness ratings was only moderate, for the same 16 cases x 12 items, with a correlation between raters of .669 ($p < .001$) and a significant difference in their average ratings (4.06 vs 3.85; $t = 3.05$, $p < .01$). Discussion of the appropriateness ratings clarified the nature of the disagreements in this measure, and the items were rescored by the same scorer.

1:10 Written English Discourse (Multiple Choice)

Inter-item reliability was investigated for the total test for the 87 students who completed the test ($\alpha = .799$).

1:11 Written English Grammar (Multiple Choice)

The written English grammar test included three subtests which differed in form. In items 1 to 15, the student was required to choose which if any of three sentences was grammatically incorrect (four choices); in items 16 to 25, to indicate which if any of three phrases contained a mistake in grammar (four choices); in items 26 to 35 (three choices), to decide whether a sentence was complete or not, and if not, which of two phrases would complete it. One of the 'wrong sentence' items was excluded from analysis because of disagreement among adult native speakers about which was the correct answer. Inter-item reliability for the total test and the three subtests is shown in Table C.10.

1:12 Written English Sociolinguistic Test

Letter and note. The letters and notes were scored by two native English speakers.

Rewriting task. The six items requiring rewriting of a phrase within a paragraph were scored by three native speakers of English. Inter-rater reliabilities were computed following clarification of the scoring criteria and were generally adequate to good.

Reliability is reported in Table C.11 for the measures of mitigation in the letter and global appropriateness in the letter and in the note, as well as for appropriateness in each of the six 'rewrite' items. In the rewriting task, the number of cases varies from item to item because of students not attempting the item or giving 'unscorable' answers.

Table C.1

Inter-rater Reliability on Expressive Measures of Oral Portuguese Grammar
(n. of cases = 21)

	Mean Difference	t	p	r	p
Finite verbs	0.24	1.16	ns	.97	<.01
Total errors	-0.38	-1.00	ns	.76	<.01
Preposition range (Story, excluding 'a')	-0.14	-0.65	ns	.75	<.01
Preposition range (Picture, excluding 'a')	-0.14	-1.00	ns	.88	<.01
Preposition count (Picture, excluding 'a')	-.05	-0.19	ns	.79	<.01

Table C.2

Inter-rater Reliability for Oral Portuguese Discourse on the Story-Telling Task
(n. of cases is 21 for detail scores; 23 for the global rating)

	Mean Difference	t	p	r	p
Logic	0.0	0.44	ns	0.87	<.01
Anaphora	0.3	1.75	ns	0.83	<.01
Time orientation	0.2	-1.14	ns	0.77	<.01
Global rating	0.05	0.19	ns	0.90	<.01

Table C.3

Inter-rater Reliability on Detail and Global Scoring of the Oral Portuguese Sociolinguistic Test
(n. of cases is based on 21 students x 12 items)

	Agreements	N of Cases	Mean Difference	t	p	r	p
Form of address	90%	174	.01	1.00	ns	.99	<.01
Use of person	85%	175	.04	1.71	ns	.95	<.01
Use of modality	89%	203	.06	2.05	<.05	.85	<.01
Extra politeness markers	93%	204	.02	1.42	ns	.98	<.01
Global rating	78%	206	.12	2.92	<.01	.79	<.01

Table C.4

Inter-item Reliability for Written Portuguese Grammar Multiple Choice
(n. of cases = 47)

	No. of Items	alpha	alpha if "worst" item deleted
Total Test	49	.88	.89
Agreement of adjectives	4	.61	.60
Articles	4	.47	.57
Prepositions	13	.71	.73
Agreement of verbs	7	.47	.50
Verb tenses	9	.43	.48
Word order	12	.59	.62

Table C.5

Inter-rater Reliability of Written Portuguese Sociolinguistic Measures
(maximum n. of cases = 21 students x 2 items = 42)

	Agreements		Reliability					
	Letter	Note	No of Cases	Mean Diff.	t	p	r	p
Opening	15	11	19	0.0	0.0	ns	1.00	<.01
Use of person	13	11	16	0.0	0.0	ns	1.00	<.01
Use of modality	9	11	26	.04	.33	ns	.76	<.01
Politeness markers	13	11	26	.15	1.44	ns	.74	<.01
Closing	15	10	15	0.0	0.0	ns	1.00	<.01
Global rating	10	4	25	.16	0.89	ns	.72	<.01

Table C.6

Inter-rater Reliability for Preposition Use in Oral English
(n. of cases = 14)

	Mean Difference	t	p	r	p
Errors	0.79	2.24	<.05	0.66	<.05
Obligatory contexts	0.21	0.76	ns	0.97	<.01
Types	0.93	2.51	<.05	0.70	<.01
Correct tokens	1.00	2.19	<.05	0.94	<.01
Errors/obl. contexts	0.09	2.57	<.05	0.65	<.05
Types/corr. tokens	0.01	0.69	ns	0.94	<.01

Table C.7

Inter-rater Reliability of English Oral Syntax and Pronunciation Ratings
(n. of cases = 78)

	Mean Difference	F	p	alpha
Oral syntax	0.00	0.0	ns	0.87
Pronunciation	0.05	1.61	ns	0.91

Table C.8

Inter-rater Reliability of Oral English Discourse Measures.

	No of Cases	Mean Difference	F	p	alpha
Global	81	0.01	0.04	ns	0.92
Scene	81	0.04	0.36	ns	0.93
Logic	78	0.03	0.18	ns	0.93
Anaphora	81	0.10	4.16	<.05	0.95
Connectors	81	0.15	6.40	<.05	0.94

Table C.9

Preliminary Inter-rater Reliability of Oral English Sociolinguistics
(n. = 16 cases x 12 items)

	Percent agreement	Chi-square	df	p
Form of address	96.9	359.7	4	<.001
Politeness markers	98.9	373.7	4	<.001
Modals	98.9	375.6	4	<.001
Conditional clauses	96.3	215.1	4	<.001
Mood	88.5	295.2	4	<.001
Explanatory information	97.9	364.7	4	<.001
Attenuation	98.9	375.3	4	<.001

Table C.10

Inter-item Reliability of Written English Grammar Multiple Choice
(n. of cases = 52)

	No. of Items	alpha	alpha if "worst" item deleted
Full test	34	.764	.780
Wrong sentence	14	.606	.636
Wrong phrase	10	.417	.539
Sentence completion	10	.617	.657

Table C.11

Inter-rater Reliability of Written English Sociolinguistic Measures

	Number of Scorers	Number of cases	Alpha
Mitigation, letter	2	41	.734
Appropriateness, letter	2	41	.854
Appropriateness, note	2	38	.764
Appropriateness, rewrite:			
Item 1	3	33	.892
Item 2	3	40	.876
Item 3	3	34	.860
Item 4	3	42	.794
Item 5	3	43	.706
Item 6	3	33	.663

APPENDIX D

Language Use Questionnaire and

Language Attitude Questionnaire for Grade 7

UQ6 ____________________

____

Name ____________________ ____________

Grade ____________________

School ____________________

Date of test ____________________

Date of birth ____________________

LANGUAGE USE QUESTIONNAIRE
FOR
GRADE 7 STUDENTS

April 1983

INSTRUCTIONS

Please use a pencil and answer the questions as carefully and correctly as possible. If you do not know the answer to a question, write "don't know". If you are not sure how to write something, please ask for help.

Some information questions should be answered by making an X in the appropriate box. Please read these questions carefully before deciding which box to check.

In some other questions you are asked to read certain statements and then circle a number between 1 and 5 which shows as closely as possible how you feel about each statement. Here are two examples:

1) I feel that the following foods taste delicious:

	Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree
Strawberries	1	2	3	4	5
Liver	1	2	3	4	5
Turnips	1	2	3	4	5
Peas	1	2	3	4	5

2) The following things are important for good health:

	Not at all	Slightly	Fairly	Very	Extremely
Food	1	2	3	4	5
Movies	1	2	3	4	5
Skiing	1	2	3	4	5

Put up your hand if you have any problems.

Please fill in the blanks or check the appropriate answer(s) if a list is provided.

PART A

					No. of Cases
1) My birthday is	____________, year	____________, month	____________ day		
2) I was born in:					185
CANADA	☐	____________, name the province	____________ name the place	35%	
MAINLAND PORTUGAL	☐	____________, name the province	____________ name the place	15%	
AZORES	☐	____________, name the island	____________ name the place	40%	
MADEIRA	☐	____________ name the place		4%	
OTHER	☐	____________, name the country	____________ name the place	7%	
3) The place I was born in is a	city ☐			60%	169
	town ☐			19%	
	village ☐			21%	
(As judged by Portuguese scorers on project staff, this is:		urban		66%	137
		rural		34%	

PART B

					No. of cases
4) When I came to Toronto I was ____________ old.					
age					117
5) Before I came to Toronto I lived in:					
CANADA	☐	____________, name the province	____________ name the place	4%	
MAINLAND PORTUGAL	☐	____________, name the province	____________ name the place	25%	
AZORES	☐	____________, name the island	____________ name the place	56%	
MADEIRA	☐	____________ name the place		6%	
OTHER	☐	____________, name the country	____________ name the place	9%	
6) This place was a	city ☐			37%	105
	town ☐			29%	
	village ☐			34%	
In the coders' judgement, this is			urban	35%	71
			rural	65%	

7) Before coming to Toronto did you live in any other place besides the one indicated above?

YES 11% NO 89% — 116

If yes, write
a) the name of the country ____________
b) the name of the province ____________
c) the name of the place ____________

			No. of cases
Breakdown of Q. 7 was:	Canada	10%	10
	Mainland Portugal	30%	
	Azores	30%	
	Madeira	-	
	Other	30%	
The places were coded as:	urban	75%	8
	rural	25%	

PART C

		YES	NO	No. of cases
8) I am now living with my	father	97%	3%	187
I am now living with my	mother	99%	1%	187
I am now living with my	brother(s)	77%	23%	167
I am now living with my	sister(s)	67%	33%	168
I am now living with my	grandparent(s)	25%	75%	151
I am now living with	other(s)	16%	84%	141

9) My father was born in:

					No. of cases
CANADA	☐	____________, name the province	____________ name the place	0%	170
MAINLAND PORTUGAL	☐	____________, name the province	____________ name the place	33%	
AZORES	☐	____________, name the island	____________ name the place	60%	
MADEIRA	☐	____________ name the place		6%	
OTHER	☐	____________, name the country	____________ name the place	1%	

				No. of cases
10) The place my father was born was a	city	☐	22%	170
	town	☐	29%	
	village	☐	48%	
Coders judged this to be	urban		28%	101
	rural		72%	

				No. of cases
11) He has attended:	No school	☐	14%	162
	Primary School (Grades 1 to 6)	☐	73%	
	Secondary School (Grades 7 to 12)	☐	11%	
	University or College	☐	2%	

No. of cases

12) What is your father's job? Describe what he does. Does he work for himself or someone else? Give as much information as you can. 139

Codable using the Occupational Classification Manual, Census of Canada, 1971

If he is out of work, what was his previous job? ______ 180

13) The language my father first learned as a child is Portuguese 99% / Other Romance 1% 172

name of language

14) My mother was born in: 172

CANADA	☐	______, name the province	______ name the place	0%	
MAINLAND PORTUGAL	☐	______, name the province	______ name the place	34%	
AZORES	☐	______, name the island	______ name the place	60%	
MADEIRA	☐	______ name the place		5%	
OTHER	☐	______, name the country	______ name the place	1%	

15) The place she was born was a

city	☐	22%	153
town	☐	28%	
village	☐	50%	

Coders judged this to be	urban	28%	101
	rural	72%	

16) She has attended

No school	☐	8%	165
Primary School (Grades 1 to 6)	☐	76%	
Secondary School (Grades 7 to 12)	☐	14%	
University or College	☐	2%	

No. of cases

17) What is your mother's job? Describe what she does. Does she work for herself or someone else? Give as much information as you can.

Codable using the Occupational Classification Manual, Census of Canada, 1971 — 138

If she is out of work, what was her previous job? ____ __________

Portuguese	99%
French	1
Non-Romance	1

18) The language my mother first learned as a child is ______________ 185
name of language

PART D

19) Try to remember all the schools you have attended and list them below. For each school specify where it was located, the number of years you attended it, and the language used for teaching most of the classes. (See summary on page 8a)

	1	2	3
School			
City or town			
Country			
Number of years			
Main language of instruction			

(If you have attended more than 3 schools, list the others on the back of this page.)

1.
L ________
T ________
U/R ________

2.
L ________
T ________
U/R ________

3.
L ________
T ________
U/R ________

	Never	Sometimes	About half the time	Often	Always	No. of responses
20) Before I went to school I spoke Portuguese	3	9	11	21	57	P 185
21) Before I went to school I spoke English	43	30	15	8	4	E 178
22) Before I went to school I spoke French	94	2	2	2	1	F 177
23) Before I went to school I spoke FRENCH please specify any other language	71	14	0	0	14	O 7

French	3%	
Portuguese	78%	37
English	3%	
Other Romance	16%	

School		1	2	3
n. of cases		182	91	35
Place coded as:	N =	182	91	35
Urban		97%	99%	100%
Rural		3%	1%	0%
Years attended:	N =	178	85	31
1		2%	35%	3%
2		4	31	81
3		8	21	16
4		7	9	
5		12	1	
6		7	1	
7		16	1	
8		44	0	
Language:	N =	184	92	36
Portuguese		4%	7%	6%
English		94	92	94
French		1	1	0
Romance		1	0	0

	Not at all	With some difficulty	Fairly easily	Quite easily	Extremely easily	No. of cases
24) Now I can understand English	0	3	4	25	68	186
25) Now I can speak English	1	2	4	22	71	185
26) Now I can read English	0	3	15	19	62	185
27) Now I can write English	1	6	7	25	61	186
28) Now I can understand Portuguese	1	13	23	35	27	186
29) Now I can speak Portuguese	0	13	22	36	29	186
30) Now I can read Portuguese	6	28	21	26	18	186
31) Now I can write Portuguese	11	33	19	21	16	186

32) Besides English and Portuguese, I also know ______________.
language

		No. of cases
French	97%	165
English	1	
Romance (other)	1	

The following includes only those specifying French:

This language I can

	Not at all	With some difficulty	Fairly easily	Quite easily	Extremely easily	No. of cases
a) understand	1	36	34	25	4	160
b) speak	3	38	34	19	6	160
c) read	4	34	29	26	6	160
d) write	6	38	32	20	4	160

On the next two pages either circle a number or check a box for each line.
If you speak mainly Portuguese or English to someone, circle a number to show how much for that line.
If you speak mainly some other language to someone, check Another language for that line.
If you never speak to someone - for example, if you do not have a sister - check Not applicable for that line.

33)

When I am talking	Portuguese all the time	Mostly Portuguese	Portuguese & English about equally	Mostly English	English all the time	Another language	Not applicable	No. of cases
a) to my father I speak	39	27	29	5	1	☐	☐	182
b) to my mother I speak	43	23	29	4	1	☐	☐	183
c) to my brother(s) I speak	2	0	9	30	59	☐	☐	128
d) to my sister(s) I speak	3	4	7	37	48	☐	☐	116
e) to my grandparents on my my father's side I speak	84	13	2	1	1	☐	☐	128
f) to my grandparents on my mother's side I speak	80	16	3	1	0	☐	☐	145
g) to my relatives that are living with me I speak	28	26	28	12	6	☐	☐	116
h) to my neighbours I speak	7	7	15	28	44	☐	☐	182

34) When I am spoken to	Portuguese all the time	Mostly Portuguese	Portuguese & English about equally	Mostly English	English all the time	Another language	Not Applicable	No. of cases
a) my father speaks to me in	61	17	21	1	0	☐	☐	181
b) my mother speaks to me in	60	27	12	1	1	☐	☐	181
c) my brother(s) speak(s) to me in	3	2	9	25	62	☐	☐	125
d) my sister(s) speak(s) to me in	5	4	5	33	53	☐	☐	112
e) my grandparents on my father's side speak to me in	93	5	1	2	0	☐	☐	125
f) my grandparents on my mother's side speak to me in	91	6	1	1	0	☐	☐	143
g) my relatives who are living with me speak to me in	39	17	28	9	6	☐	☐	110
h) my neighbours speak to me in	7	5	18	22	48	☐	☐	180

35) My parents usually speak to each other in Portuguese = 100% . 182
language(s) (First listed)

36) My friends and I usually speak to each other in English 98% Portuguese 2% . 183
language(s)

No. of cases

PART E

37) I watch Portuguese T.V. programs for about mean: 2.32 s.d. 3.86 hours a week. 178

no. of hours

38) I watch English T.V. programs for about mean: 22.81 s.d. 18.54 hours a week. 175

no. of hours

(For each of the following questions check one box only.)

39) In Portuguese, I read (not counting textbooks

3 or more books a week	☐	4%
1-2 books a week	☐	10%
2-3 books a month	☐	7%
1 book or less a month	☐	15%
hardly any books	☐	64%

180

40) In English, I read (not counting textbooks)

3 or more books a week	☐	35%
1-2 books a week	☐	27%
2-3 books a month	☐	24%
1 book or less a month	☐	8%
hardly any books	☐	5%

184

41) I listen to Portuguese radio programs:

every day	☐	10%
almost every day	☐	9%
weekends only	☐	14%
once in a while	☐	36%
never	☐	31%

185

42) I listen to English radio programs:

		No. of cases
4 hours or more every day	☐ 34%	184
2-3 hours a day	☐ 19%	
about 1 hour a day	☐ 22%	
less than 1 hour a day	☐ 18%	
hardly ever	☐ 7%	

43) I write letters in Portuguese to my friends and relatives:

about once a week	☐ 5%	183
about once a month	☐ 19%	
about once a year	☐ 9%	
hardly ever	☐ 21%	
never	☐ 45%	

44) I write letters in English to my friends and relatives:

2-3 times a month	☐ 16%	182
once or less a month	☐ 21%	
twice a year	☐ 12%	
once a year	☐ 20%	
never	☐ 30%	

45) Do you attend Portuguese classes in any of the Portuguese clubs?

Yes ☐21% No ☐79% 182

If yes, how many hours a week? 1-5: 31% 10-13: 69% 35

46) For how many years have you been studying in the Heritage Program? 0: 6% 1-7: 83% 8-10: 11% 170
no. of years

47) For how many years have you studied Portuguese outside the Heritage Program? 0: 59% 1-7: 39% 9-13: 2% 152
no. of years

			No. of cases
48) Are you involved in any Portuguese folk dance group?			
Yes ☐ 16% No ☐ 84%			180
If yes, a) what is the language used by the dancers?			
Portuguese all the time	☐	14%	
Mostly Portuguese	☐	14%	
Portuguese and English about equally	☐	36%	28
Mostly English	☐	21%	
English all the time	☐	14%	
b) what is the language used by the instructors?			
Portuguese all the time	☐	33%	
Mostly Portuguese	☐	22%	
Portuguese and English about equally	☐	37%	27
Mostly English	☐	4%	
English all the time	☐	4%	

In the questions below, either circle a number or check a box as you did for questions 33 and 34.

		Portuguese all the time	Mostly Portuguese	Portuguese & English about equally	Mostly English	English all the time	Another language	Not applicable	No. of cases
49)	When I go to church I attend mass in	42	19	19	12	8	☐	☐	176
50)	In other activities of the Portuguese Community (eg. special parties or special festivities) I speak	11	20	45	16	14	☐	☐	173
51)	People talking to me on these occasions speak	13	22	44	15	6	☐	☐	175
52)	When I am at school I speak	0	2	4	27	66	☐	☐	179
53)	The other students at school speak to me in	0	2	2	20	76	☐	☐	177
54)	During sports activities I speak	0	2	2	18	78	☐	☐	176
55)	People playing with me speak	1	2	3	18	76	☐	☐	178

56) In the last 5 years I have been to Portugal ____ times. — No. of cases: 162

0	50%
1	25%
2	11%
3	7%
4	3%
5	2%
6	1%

AQ6 ____________________

____

Name: ____________________

Grade: ____________________

School: ____________________

Date of test: ____________________

Date of birth: ____________________

LANGUAGE ATTITUDE QUESTIONNAIRE

FOR

GRADE 7 STUDENTS

April 1983

INSTRUCTIONS

Please use a pencil and answer the questions as carefully and correctly as possible. If you do not know the answer to a question, write "don't know". If you are not sure how to write something, please ask for help.

Some information questions should be answered by making an X in the appropriate box. Please read these questions carefully before deciding which box to check.

In some other questions you are asked to read certain statements and then circle a number between 1 and 5 which shows as closely as possible how you feel about each statement. Here are two examples:

1) I feel that the following foods taste delicious:

	Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree
Strawberries	1	2	3	4	5
Liver	1	2	3	4	5
Turnips	1	2	3	4	5
Peas	1	2	3	4	5

2) The following things are important for good health:

	Not at all	Slightly	Fairly	Very	Extremely
Food	1	2	3	4	5
Movies	1	2	3	4	5
Skiing	1	2	3	4	5

Put up your hand if you have any problems.

PART A

I feel that the following languages are important for me:

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
1) to do well in school.						
Portuguese	3	7	34	30	26	191
English	0	1	3	15	81	191
French	10	10	27	29	24	197
2) to have an important position in the Portuguese Community.						
English	2	15	13	35	36	185
Portuguese	3	4	12	19	63	188
3) to communicate with my family.						
English	8	17	24	30	21	178
Portuguese	1	3	6	17	74	191
4) to get into a good university.						
Portuguese	5	9	30	20	36	185
English	1	1	2	10	86	191
French	5	7	11	33	44	184

I feel that the following languages are important for me:

5) to make good friends.

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
English	2	1	7	17	73	190
Portuguese	5	11	34	30	19	185

6) to get into a good high-school.

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
Portuguese	5	17	27	32	19	186
English	1	1	3	11	84	192
French	5	4	18	40	32	185

7) to communicate with my friends.

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
English	1	1	4	16	79	191
Portuguese	6	13	33	34	13	186

8) to get a good job in the future.

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
Portuguese	3	8	17	24	47	188
English	0	0	2	8	90	192
French	6	7	11	30	45	187

I feel that the following languages are important for me:

9) to be more comfortable with people around me.

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
Portuguese	1	10	24	35	30	187
English	1	0	4	17	78	189

10) to be a well educated person.

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
English	2	1	3	10	85	189
Portuguese	1	6	14	26	53	186
French	5	3	11	25	56	187

PART B

11) I feel comfortable speaking:

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
Portuguese	3	12	32	35	18	186
English	1	1	2	13	83	192
French	16	26	33	16	9	185

12) I like:

English	0	1	1	13	85	189
Portuguese	3	5	19	37	38	189
French	19	16	24	26	15	186

13) I would like to be taught in:

English	0	0	3	9	89	188
Portuguese	4	10	22	28	36	186
French	17	18	17	29	19	185

14) I like to watch television in:

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
Portuguese	17	22	33	16	12	186
English	1	1	1	8	90	191
French	47	27	17	6	3	186

	Portuguese all the time	Mostly Portuguese	Portuguese and English about equally	Mostly English	English all the time	
15) To talk about my feelings I like to use:	3	4	24	30	40	189
16) To talk with my friends I like to use:	1	1	6	26	66	189
17) When I talk to my family I like to use:	30	26	35	7	2	191
18) When my family talks to me I like them to use:	27	22	36	8	6	191
19) I like to talk about things that happen in school in:	6	11	21	24	39	191

PART C

	(check one only)	No. of responses
20) I consider myself to be mainly:		
a) a Portuguese speaker	☐ 5%	
b) an English speaker	☐ 10%	
c) a speaker of English, French and Portuguese	☐ 30%	
d) a speaker of English and Portuguese	☐ 55%	192
21) When I grow up I would like to be considered as:	(check one only)	
a) a speaker of English, French and Portuguese	☐ 52%	
b) a Portuguese speaker	☐ 2%	
c) a speaker of English and Portuguese	☐ 40%	
d) an English speaker	☐ 5%	191
22) When I grow up I would like to marry:	(check one only)	
a) a speaker of English and Portuguese	☐ 65%	
b) a speaker of English, French and Portuguese	☐ 23%	
c) an English speaker	☐ 7%	
d) a Portuguese speaker	☐ 5%	191
23) If I have children, I would like them to be:	(check one only)	
a) English speakers	☐ 5%	
b) speakers of English and Portuguese	☐ 58%	
c) Portuguese speakers	☐ 4%	
d) speakers of English, French and Portuguese	☐ 34%	191

PART D

Please give your opinion on the following statements:

24) The Portuguese language should be kept alive in Toronto.

	Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree	No. of responses
	1	2	15	38	45	191

25) The Portuguese Community would disappear if all the Portuguese people knew English very well.

	Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree	No. of responses
	13	17	26	26	17	189

26) In Toronto people recognize the value of the following languages:

	Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree	No. of responses
English	0	0	3	17	80	189
Portuguese	1	2	22	51	25	186
French	7	10	29	35	20	185

27) In the world people recognize the value of the following languages:

	Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree	No. of responses
Portuguese	2	7	26	34	32	183
English	0	0	5	20	74	186
French	6	7	23	33	32	183

28) English-Canadians think it is very important for a person to study the following languages:

	Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree	No. of responses
Portuguese	3	6	27	38	26	187
English	1	0	2	14	83	191
French	5	5	13	37	40	188

29) In Toronto people recognize the value of the following groups:

Portuguese-Canadians	1	4	20	45	30	186
English-Canadians	0	2	7	21	70	188
French-Canadians	6	8	17	37	32	185

30) The following groups are very proud of their history, language and culture:

English-Canadians	1	1	4	26	69	188
Portuguese-Canadians	1	3	9	33	55	188
French-Canadians	4	3	19	28	46	186

31) People in the Portuguese Community in Toronto have a lot of contact with the following groups:

English-Canadians	0	4	14	36	45	183
Portuguese-Canadians	1	2	8	31	59	186
French-Canadians	13	14	30	28	15	184

PART E

		Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree	No. of responses
32)	In my home people speak very good Portuguese.	0	2	12	36	51	192
33)	In my home people speak very good English.	7	10	38	32	12	191
34)	The Portuguese spoken in the Azores is as good as the Portuguese spoken in the Mainland.	10	19	25	29	16	192
35)	An Azorean student can have problems understanding a Portuguese teacher from the Mainland.	6	13	27	38	17	191
36)	A student from the Mainland can have problems understanding a Portuguese teacher from the Azores.	6	12	28	37	18	191
37)	When speaking English it is all right to use some Portuguese words.	24	21	28	16	11	191

38) When speaking Portuguese it is all right to use some English words.

Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree	No. of responses
16	16	25	31	12	189

39) I think that in the Portuguese classes we should be taught in:

	(check one only)		No. of responses
Azorean-Portuguese	☐	15%	192
Mainland-Portuguese	☐	23%	
Both Azorean and Mainland Portuguese	☐	21%	
It does not matter	☐	41%	

PART F

How well do the following statements reflect your parents' views?

40) My parents would like me to go to university.

Not at all	Slightly	Fairly	Very	Extremely	No. of responses
2	3	10	26	60	191

41) My parents feel that a higher education is more important than getting a job when I reach 16.

4	4	8	17	67	192

42) My parents like to speak with my teachers to find out how I am doing in school.

1	3	11	21	64	192

43) My parents feel it is important for me to study the following languages:

Portuguese	1	1	3	20	76	192
English	0	1	3	13	84	189
French	5	8	19	34	34	187

44) My parents feel that a Portuguese-Canadian should know the following languages very well:

	Not at all	Slightly	Fairly	Very	Extremely	No. of responses
English	0	2	7	20	72	190
Portuguese	1	0	3	23	73	188
French	8	13	28	33	18	187

45) When I have homework in the following languages my parents make sure I do it.

	Never	Sometimes	About half the time	Often	Always	
Portuguese	6	9	8	20	56	186
English	2	2	5	9	83	186
French	13	17	16	21	33	182

PART G

46) I give my opinion on subjects when I am speaking with

	Never	Sometimes	About half the time	Often	Always	No. of responses
a) adults outside my family	6	22	19	32	21	190
b) my parents	2	7	13	36	43	191
c) my brother(s) and sister(s)	5	5	11	26	52	183
d) my teachers	4	15	21	32	28	191
e) my friends	1	6	15	32	46	191

47) In the future I am planning: (check one only)

a) to stay in Canada	☐	87%	184
b) to go back to Portugal permanently	☐	13%	

48) My family is planning: (check one only)

a) to stay in Canada	☐	73%	182
b) to go back to Portugal permanently	☐	27%	

49) If I could, I would prefer to spend my holidays in: (check one only)

a) Canada	☐	4%	
b) Portugal	☐	35%	188
c) Other countries	☐	32%	
d) all of the above	☐	29%	

50) I know Portuguese history and achievements.

Not at all	Slightly	Fairly Well	Very Well	Extremely well	No. of responses
8	24	41	19	8	189

51) I know Canadian history and achievements.

3	9	29	42	17	189

52) I feel proud of Portuguese history and achievements.

Strongly Disagree	Disagree	Mixed Feelings	Agree	Strongly Agree	
0	5	28	34	33	189

53) I feel proud of Canadian history and achievements.

2	3	21	32	42	189

Thank you for helping us with this project.

Before you hand in this questionnaire, please go back and make sure that you have answered every question as well as you can.

If you do not know the answer to a question, please write "don't know".

If you are not sure how to say something, please ask for help.

APPENDIX E

TABLE E.1

Number of Cases, Mean and Standard Deviation for Selected Predictor Variables amd Measures of Portuguese Proficiency for the Sample Excluding Outliers on Proficiency Measures and Students with less than 3 Years Residence

Variable	Mean	SD	Cases
FSJOB	12.4510	2.7735	51
PFAMUSE4	.5606	.5001	66
TALKWPAR	1.8346	.7002	66
TALKWSIB	4.4955	.8242	55
MEDIAP1	2.0769	.9660	65
TVHRSP	1.7538	2.2152	65
VISITS	1.0169	1.2525	59
MASS	2.1111	1.4155	63
COMMEVNT	3.0476	1.0988	63
CLASSA	1.0690	1.3874	58
INSTRP2	3.8545	.8151	63
VITALP2	3.8170	.5232	64
VITALE1	4.7313	.3940	64
VITALF1	3.9156	.8922	64
ACCPTP2	3.8594	.8008	64
ACCPTE2	4.8256	.3958	65
ACCPTF2	2.9590	1.0713	65
IMAKFRNP	3.4603	1.1191	63
ICOMFAMP	4.5758	.7658	66
ICOMFRNP	3.2727	1.1031	66
ICOMPEOP	3.8333	1.0013	66
ICOMFAME	3.3651	1.2353	63
LIKETVP	2.6364	1.1979	66
LIKEFRND	4.5606	.7260	66
LIKEFAM	2.2537	1.0201	67
LMIXG	3.0000	1.0753	65
GOODHOMP	4.2090	.7290	67
EDMITJOB	4.5970	.7989	67
HOMWORKP	4.0313	1.3330	64
KANDPP	3.4403	.8002	67
KANDPE	3.8134	.7326	67
ORIENTPX	2.5667	.7449	60
AOA	2.5373	2.5425	67
POGITOTL	.8462	.0898	65
POGIGRAM	.8387	.1122	65
POGIPHON	.9607	.0772	65
POGIVOCA	.7973	.1286	65
POGERRAT	.3398	.1946	62
PODGLOBL	3.0000	1.1592	65
PWDMULTC	.5118	.1640	65
POSDIFF4	1.0576	.2975	61
POSAVAPP	4.4516	.2951	64
PORTORAL	3.7836	.9055	67
PORTWRIT	3.1493	1.1045	67

TABLE E.2

Partial Correlations between Measures of Portuguese Proficiency and Selected Predictor Variables Controlling for Age on Arrival
for the Sample Excluding Outliers and Students with less than 3 Years Residence

	POGITOTL	POGIGRAM	POGIPHON	POGIVOCA	POGERRAT	PODGLOBL	PWDMULTC	POSDIFF4	POSAVAPP	PORTORAL	PORTWRIT
FSJOB	-.1704	-.0722	-.1358	-.2324	.1601	-.3487*	-.2643	-.3343	-.2794	.1159	.1550
PFAMUSE4	-.0698	-.0478	-.1247	-.0445	-.2514	-.0065	.0590	.1024	.0855	.2609	.0094
TALKWPAR	.1209	.0875	.1783	.0832	.1610	.0346	.0926	.1079	-.0549	-.2754	-.0266
TALKWSIB	-.2537	-.1968	-.0732	-.2443	.2016	-.3017	-.2031	-.1085	-.1805	.0663	-.0191
MEDIAP1	.3923**	.3115*	.1039	.3752*	-.3230*	.2788	.2042	-.0195	-.0410	.5535**	.5959**
TVHRSP	.2857	.2395	.0832	.2563	-.2634	.2441	.3254*	.1811	.2607	.3718*	.4423**
VISITS	.4860**	.4382**	.2023	.3987*	-.2488	.3359*	.1380	.1286	.1233	.4108**	.3046
MASS	-.3047*	-.3481*	-.0054	-.1861	.1390	-.2312	-.2127	.0549	.0584	-.3367*	-.2933
COMMEVNT	-.1932	-.2689	-.0674	-.0331	.1536	-.2510	-.1499	-.0036	-.0618	-.0127	-.0121
CLASSA	.4051*	.4536**	.0460	.2289	-.2995	.3814*	.4892**	-.0470	.0215	.3545*	.4348**
INSTRP2	.0214	.1168	-.0935	-.0813	-.1928	.0317	.2035	-.0353	-.0453	.2921	.3143*
VITALP2	-.1876	-.0788	-.2000	-.2334	-.0522	-.2069	-.0744	-.0228	-.1212	.1011	.0208
VITALE1	-.0878	.0383	-.0605	-.2176	.0412	.0277	.1977	.0848	.1084	.1150	.0495
VITALF1	.2864	.3592*	.0622	.1247	-.3364*	.4006**	.2605	.1471	.2504	.1366	.1244
ACCPTP2	.1107	.1185	-.0539	.0920	-.1250	.1125	.1254	.1763	.2099	.4686**	.4291**
ACCPTE2	-.1428	-.0911	-.2418	-.0997	.0649	.1340	.1245	.1430	.1637	.1154	.0488
ACCPTF2	.1569	.1754	-.1059	.1308	-.2086	.1451	.2664	.0370	.0083	.2929*	.5217**
IMAKFRNP	.0281	.1658	-.1885	-.0999	-.0964	.0335	.0083	.0086	.0098	.2328	.2531
ICOMFAMP	.0249	.0235	.0355	.0102	-.1123	.1433	.2522	.1223	.2068	.3042*	.1932
ICOMFRNP	.1027	.1726	-.0477	.0016	-.2819	.0774	.0053	.1101	.1698	.2629	.1822
ICOMPEOP	.1024	.0789	.0527	.0960	-.1786	.0943	.1293	.2061	.2111	.1198	.0222
ICOMFAME	.1297	.1992	-.1326	.0407	-.1473	.1161	.2389	.0679	-.0980	.3209*	.3754*
LIKETVP	.2187	.2742	-.0011	.0980	-.0402	.1690	.3543*	.0015	.1413	.3039*	.4945**
LIKEFRND	.0351	.1324	.0451	-.1192	-.0708	.1300	.1253	.0171	.0271	.0588	-.0013
LIKEFAM	.0355	.0775	-.0432	-.0152	-.0073	-.0937	.0587	.0622	-.0838	-.2558	-.2605
LMIXG	-.2541	-.2812	-.0973	-.1322	.2307	-.3864**	-.1376	-.2098	-.2155	-.2729	-.1157
GOODHOMP	.1151	.1722	-.0229	.0207	.0674	.2021	.1672	.1024	.1133	.2654	.2435
EDMITJOB	-.0035	.0367	-.0454	-.0428	.0938	.1532	.2543	.1505	.1532	.2112	.0849
HOMWORKP	.0173	-.0735	.0588	.1192	-.0767	.1984	.1727	-.0194	.1321	.2008	.3313*
KANDPP	.3169*	.2163	.1569	.3370*	-.1784	.2622	.3218*	.1832	.1002	.3887**	.5524**
KANDPE	-.0600	-.0902	-.1069	.0278	-.0335	.1159	.1288	.0055	-.0161	.2208	.3419*
ORIENTPX	.2389	.2311	.0724	.1754	-.1587	.2339	-.1567	-.1059	-.0636	.3362*	.0514

* - SIGNIF. LE .01
** - SIGNIF. LE .001

TABLE E.3

Number of Cases, Mean and Standard Deviation for Selected Predictor Variables and Measures of English Proficiency for the Sample Excluding Students with less than 3 Years Residence

Variable	Mean	SD	Cases
FSJOB	12.6545	2.7703	55
PFAMUSE4	.5634	.4995	71
TALKWPAR	1.8638	.7088	71
TALKWSIB	4.5297	.8055	59
MEDIAP1	2.0290	.9595	69
TVHRSP	1.6429	2.1739	70
VISITS	1.0938	1.3884	64
MASS	2.2206	1.4645	68
COMMEVNT	3.1045	1.1165	67
CLASSA	1.0000	1.3560	63
INSTRP2	3.8234	.8292	67
VITALP2	3.7785	.5627	69
VITALE1	4.7304	.3997	69
VITALF1	3.9159	.8714	69
ACCPTP2	3.7874	.8343	69
ACCPTE2	4.8238	.3961	70
ACCPTF2	3.0095	1.0752	70
IMAKFRNP	3.3676	1.1448	68
ICOMFAMP	4.5493	.7890	71
ICOMFRNP	3.2113	1.1201	71
ICOMPEOP	3.7465	1.0243	71
ICOMFAME	3.3235	1.2025	68
LIKETVP	2.5634	1.2040	71
LIKEFRND	4.5915	.7087	71
LIKEFAM	2.2917	.9991	72
LMIXG	2.9714	1.0864	70
GOODHOMP	4.1944	.7246	72
EDMITJOB	4.5417	.8381	72
HOMWORKP	4.0580	1.3048	69
KANDPP	3.4375	.8090	72
KANDPE	3.8264	.7607	72
ORIENTPX	2.5385	.7305	65
AOA	2.4167	2.5105	72
EPRPERAT	.1758	.1362	69
ECUNITSZ	8.9952	1.5839	72
EODDETL4	3.2257	.9252	72
EWDMULTC	.6325	.1462	71
EOSDIFF	.2634	.1159	72
EOSAVAPP	4.1505	.3413	72
ENGLORAL	4.6319	.6109	72
ENGLWRIT	4.4097	.7887	72
FRENCH	2.9462	.8309	65

TABLE E.4

Partial Correlations between Measures of English Proficiency and Selected Predictor Variables Controlling for Age on Arrival for the Sample Excluding Students with less than 3 Years Residence

	EPRPERAT	ECUNITSZ	EODDETL4	EWDMULTC	EOSDIFF	EOSAVAPP	ENGLORAL	ENGLWRIT	FRENCH
FSJOB	.0174	-.3323*	-.2563	-.4366**	-.1160	-.0086	-.0981	.0271	.2196
PFAMUSE4	-.0537	.0561	-.2048	.0342	-.0853	.1165	.0250	-.0497	-.0761
TALKWPAR	-.1339	-.0534	-.0486	.0110	.1257	-.0517	.0883	.1842	.0822
TALKWSIB	-.0258	-.3056*	-.3702*	-.1167	-.2345	.0115	.1979	-.0493	-.0687
MEDIAP1	.2984*	-.1839	-.0390	.0311	.0160	.0707	.0475	.1430	.3548*
TVHRSP	.1122	-.0518	.1073	.1705	-.0958	.2374	.0277	.1775	.3942**
VISITS	.0572	.0013	-.0046	.0335	-.1234	-.1689	.1723	.1127	.2186
MASS	-.2898	-.2010	-.0429	-.2934*	.0199	.0121	-.0544	-.1752	-.2571
COMMEVNT	-.0210	.1666	-.1658	-.2954*	-.1485	-.2223	.0526	.0025	.0708
CLASSA	.1815	.1813	.0891	.1389	-.0575	-.1479	.0606	.0311	.1988
INSTRP2	.2058	-.1130	.0264	.0815	.1169	.0800	.1635	.2098	.2508
VITALP2	.0984	-.1043	-.0614	.0070	.0917	.0958	.0939	.0876	.0254
VITALE1	-.1070	-.0195	.1284	.1722	.1037	.3011*	.3611*	-.0114	.0129
VITALF1	.1133	.1631	.1217	.4337**	.0405	.2404	.1548	.2413	.1190
ACCPTP2	.2425	-.0889	.0525	.0504	.0585	.1642	.1309	.1801	.2682
ACCPTE2	-.0801	.2173	.0285	.1649	.1634	.0464	.5167**	.1685	-.0114
ACCPTF2	.2338	.0421	.0562	.1211	-.0653	.0752	.1370	.4245**	.6706**
IMAKFRNP	.0794	-.0359	.0285	-.0865	.0768	-.0837	-.0808	.0580	.1806
ICOMFAMP	.2159	.0209	-.0840	.3700**	.0458	.2656	.2413	.3905**	.0080
ICOMFRNP	.0506	-.1374	-.0019	.0307	.0737	.0724	.0075	.0370	.2435
ICOMPEOP	-.0650	.0601	-.0481	.2840*	.2507	.3404*	.2476	.3868**	.0803
ICOMFAME	.0861	.0705	.2992*	.1300	-.0260	.0198	.1911	.1319	.4293**
LIKETVP	.2344	.0977	.0944	.0943	-.3208*	-.0654	.0673	.1788	.3524*
LIKEFRND	.1055	-.0450	.0191	.2764	-.0604	.1549	.2576	.0549	.2076
LIKEFAM	-.0211	.0709	.1399	.1617	.1776	.0606	.1116	-.0465	-.0560
LMIXG	.0930	-.0610	-.2276	-.3264*	.0559	-.1273	.0150	-.0966	-.1921
GOODHOMP	-.0893	-.2312	.0488	.1087	-.0703	-.0538	.3152*	.0662	.0121
EDMITJOB	-.0452	-.0205	-.0367	.0556	.2158	.2296	.2977*	.0906	.0911
HOMWORKP	.0972	-.1783	.0303	.0913	-.1316	.0207	.1591	.1553	.1736
KANDPP	.1175	-.0606	.2177	.0798	-.0772	-.0175	.2237	.2280	.4873**
KANDPE	.1182	-.1183	.0225	.0551	.0608	-.0055	.3258*	.2639	.3908**
ORIENTPX	.0353	-.2284	-.1645	.0463	-.1078	-.1843	-.1564	-.0764	-.0867

* - SIGNIF. LE .01
** - SIGNIF. LE .001

Chapter 8

THE DEVELOPMENT OF BILINGUAL PROFICIENCY IN THE TRANSITION FROM HOME TO SCHOOL: A LONGITUDINAL STUDY OF PORTUGUESE-SPEAKING CHILDREN

Jim Cummins, José Lopes and Jorge Ramos

1. PURPOSE OF THE STUDY

This ongoing longitudinal study addresses issues concerned with the development of literacy among children of Portuguese background. The major purpose is to study the development of bilingual proficiency among school-aged children from junior kindergarten (JK) to grade 1. The study's primary goals are to:

(1) describe the processes through which Portuguese-background children in the early years of school develop proficiency in both English and Portuguese conversational and academic skills;

(2) compare the development of children's Portuguese language and literacy-related skills development with that of Portuguese children in the Azores, for whom similar data have also been collected.

(3) relate patterns of interaction in children's homes both to their current skills in Portuguese and English and also to their subsequent academic progress in each language;

(4) investigate the relationships between the development of proficiency in English and the proficiency levels that children have attained in Portuguese;

(5) assess the extent to which the L1 and L2 language and literacy-related measures that we have identified are sensitive to potential learning problems that children from minority language backgrounds may encounter; such problems are often difficult to identify at an early stage in the child's schooling because the language difference masks the learning problem.

2. METHODOLOGY

2:1 Subjects

The sample for the longitudinal study consists of 20 students enrolled in Portuguese heritage language classes who in 1985 were attending JK in a school of the Metropolitan Separate School Board in the City of Toronto. During JK and senior kindergarten (SK) each one of these students attended a 30-minute daily program of Portuguese language and culture that is integrated in the school day. The school, where about 90% of the students are of Portuguese background, is located in a predominantly Portuguese-populated area in the inner city. The students were all born in 1980 and were five years old at the time of initial data collection.

Participation in the study was voluntary and dependent on the students obtaining letters of permission from parents. Of all the permission letters sent out to parents in the two JK classes at the school, 26 were returned with parents' agreement; of these, six subsequently refused to permit taping of their children in the home setting. Therefore, the final sample for the study consisted of 20 students.

In addition to this main sample, data were collected from children in the Azores and Mainland Portugal for comparison purposes. Over a period of three weeks (April and May 1984), tests were administered to 26 grade 1 students in Ponta Delgada (São Miguel, Azores). The students had permission from the parents and were in two different classes. The measures administered at the school were as follows: (1) Concepts About Print (Clay 1981), (2) Interview and (3) Test of Writing Vocabulary (Clay 1981). School records were made available to the research officer in charge of testing and information was gathered on children's socioeconomic status, based on parent occupation. A large majority of these students come from a background similar to those in Toronto. Due to the fact that in Portugal the average age for grade 1 students is around seven, we chose, in addition, six five-year-old students, whose parents permitted them to participate in the study, to be taped for one and a half hours for a recording of spontaneous oral language. Of these six children, according to the information collected on parents' education and occupation, four were of middle class urban background. These were taped in the home after 5 p.m. The remaining two were taped during the day at two state kindergartens located in rural areas of the Island of São Miguel. These students were of rural background. All the students that make up the Azorean sample were exposed to foreign languages only through television and radio.

In addition to the native speaker data collected in the Azores, data were also collected in Mainland Portugal. These data involved ten five-year-old children from a village situated 100 kilometres northwest of Lisbon. All the children had parents' permission to participate in the study. A Portuguese version of the Sentence Repetition Test, the Record of Oral Language (ROL) (Clay et al 1976), was administered to the ten children. Out of a sample of ten, six children were randomly chosen to be taped in the home. All tapings in the home took place after 5 p.m. Notes on the parents' professional and educational background were taken, as well as on the children's and parents' place of birth. They were also asked questions related to the amount of possible exposure to foreign languages. The children were only exposed to foreign languages through television and radio and were all of rural background with parents born in the same village.

2:2 Instruments

The main Toronto sample of 20 students is being followed over a period of three years: Year 1 — JK, Year 2 -- SK, Year 3 — grade 1. The following assessment procedures have been administered:

Year 1:

School (June 1985)	Home (June to July 1985)
(a) Concepts about Print (Clay 1981) (English and Portuguese)	(a) Tapings in the home
(b) Sentence Repetition (Clay et al 1976) (English and Portuguese)	(b) Parents' Questionnaire
(c) Letter Identification (Clay 1981) (English and Portuguese)	
(d) Draw a Person Test (Harris 1963)	

Year 2:

School (May to June 1986)	Home (May to July 1986)
Same tests as in Year 1 above	Tapings in the home

Year 3:

School (May-June 1987: projected date)	Home
(a) Writing Vocabulary Test (Clay 1981) (English and Portuguese)	Tapings in the home
(b) Conversation Task	
(c) Gates-McGinitie Reading Tests (Reading Subtest)	

All the testing pertaining to Years 1 and 2 has been completed. The Year 3 testing will take place in May-June 1987.

The students were taped in the home for a one and a half hour period in JK and SK. The tapings took place around the end of the school year and after 5 p.m. in order to permit analysis of the interaction between the children and the parents. In the first year of the study a questionnaire on language use and attitudes was developed and used in an interview situation with the parents. The questionnaire was filled in by one researcher while a second researcher wrote down any pertinent comments or opinions offered by the parents about their children and about the heritage language program. The various instruments administered in the school to the children are described below.

1. **Concepts About Print.** This test was developed by Marie Clay and appears in her book The Early Detection of Reading Difficulties (Clay 1981). The purpose of this test is to assess children's knowledge about the conventions of written language and the children's experience of literacy prior to school. In this test the children are shown a booklet with a story. The booklet contains text as well as illustrations. The children are told from the beginning to concentrate on the text rather than on the illustrations. The

tester then reads to the child the text on each page and asks him/her a set of questions pertaining to certain conventions of the written language used in the text. For example, children are asked where one should start reading, or for more advanced items, they might be asked to point to the misspelled word on the page. The test was administered individually to each child in a quiet room and was scored by the tester during administration. The scoring was done according to the scoring criteria in the manual, i.e., each question out of a set of 26 was scored as either right or wrong. For the English version of this test the booklet Stones was used. The Portuguese test consists of a translation of the booklet Sand (Areia) and of essentially the same set of questions; however, two questions were added concerning the use of diacritics in Portuguese. The test was administered to all students in the main sample in Years 1 and 2. The Portuguese version of the text was also administered to the grade 1 Azorean students in the comparison group.

2. **Sentence Repetition.** The Record of Oral Language (Clay et al. 1976) was used to provide both a quantitative index of children's level of language development and qualitative information about which grammatical aspects of the language (in both English and Portuguese) children had and had not yet mastered. The test consists of 42 sentences grouped into three levels of increasing complexity, level 1 being the easiest level. At each level there were seven types of sentences, as follows:

Type A: SUBJECT + VERB 'TO BE' + SIMPLE STATEMENT

Type B: SUBJECT + VERB + DIRECT OBJECT

Type C: SUBJECT + VERB + ADDITIONAL CONSTRUCTION (which should not be a direct object)

Type D: SUBJECT + VERB + INDIRECT OBJECT + DIRECT OBJECT

Type E: SUBJECT + VERB + NOUN CLAUSE

Type F: ADVERB OR RELATIVE PRONOUN + VERB + SUBJECT

Type G: SUBJECT + VERB PHRASE + OBJECT + ADDITIONAL CONSTRUCTION (e.g., adverb or adverbial phrase).

For example, sentences of Type E were as follows in the different levels: Level 1 — I guess we're lost."; Level 2 — "The boy saw what the man was doing to the car."; Level 3 -- "The teacher knows how much wood we will need for the house". These sentences were read individually to each child starting with Level 1 and the child was asked to repeat them; however, if the students had difficulty with three or more sequential items at a certain level, testing was terminated. Prior to the 1986 data collection, some vocabulary items in the English version of this test were changed due to the fact that it was felt they were not commonly used in Canadian English (note that this test was initially developed for New Zealanders); consequently, eight sentences were altered. The following are some of these changes: (a) "There is my baby riding in his pushchair" changed to "There is my baby riding in his carriage"; (b) "The shopkeeper sold my mummy some fresh cream" changed to "The man sold my mummy some fresh milk" and (c) That dog and the one next door like to chase the postman" changed to "That dog and the one next door like to chase the mailman". A Portuguese version of this test was developed, also grouping into three levels of complexity sentences that had the same syntactic features as the English ones. Each student's performance in this test was taped, transcribed by native speakers of each of the two languages, and checked by other native speakers prior to scoring. Each one of the items in the test was scored for: (a) exact repetition, i.e., use of the same words in the same order excluding dialectal variants, (b) exact repetition allowing for the use of dialectal variants affecting, in this case, certain pronouns, prepositions and verb endings, (c) exact meaning, i.e., maintenance of the original meaning in spite of possible changes that might include grammatical errors, and/or lexical additions, omissions and substitutions, and (d) number of grammatical errors and/or mistakes consisting of addition, omission or substitution of any word. Note that mistakes concerning word order were not taken into account in (c) and (d).

This test was administered to all students in the main sample in Years 1 and 2. The Portuguese version of the test was administered to the ten children that constituted the Mainland Portugal native speaker sample.

3. **Letter Identification.** This test is described in Clay (1981) and was administered and scored according to the prescribed procedures, in both English and Portuguese. The purpose of the test is twofold: (1) to examine how familiar the children were with the names of the letters of the alphabet in both English and Portuguese and (2) to provide an additional index of children's preschool knowledge of literacy. In the Letter Identification Test the children were shown a sheet of paper with all the alphabet letters. Prior to the test, the children were reminded of the language to be used (i.e., either English or Portuguese) while identifying the letters. The letters were shown both in capital and small letters. The testers pointed to a letter and asked the children to identify it. The children were tested individually and their responses recorded on tape. The scoring was done as the test progressed and a tape-recording was used for reliability purposes. The children's responses were scored either as right or wrong.

This test was administered to the children in the main sample only in Years 1 and 2.

4. **Draw a Person Test.** This test was adapted from the Draw a Man and the Draw a Woman Tests (Harris 1963). It is designed as a measure of the children's intellectual maturity. Unlike the detailed scoring proposed by Harris, which varies according to the gender of the person drawn, the drawings are being scored globally on a five-point scale taking into account, where appropriate for the age group concerned, the detailed scoring criteria prescribed in the administration manual (e.g., presence of features such as head, eyes, fingers, clothing articles, etc...). This test was administered in a group situation to the children; the two testers involved gave the instructions in both English and Portuguese and the children then proceeded to draw a person. Pencils and a blank sheet of paper were provided by the testers.

This test was administered to the children in the main sample only in Years 1 and 2.

5. **Tapings in the home.** The spontaneous oral language of children was recorded by means of an AIWA (Model TP-M6) voice-activated micro-cassette recorder which was worn by the children. The purpose of this recording was to analyse the level of acquisition of grammatical and pragmatic features of both English and Portuguese, as well as possible patterns of code switching. All children in the main sample were taped

in Years 1 and 2 after 5 p.m. for a one and a half hour period in the home. During Year 1 each family was visited initially by a male and a female researcher but during Year 2 the data were collected by just one researcher. Children were encouraged to speak whatever language they wanted (English or Portuguese). They were also told they should go about their normal routines as usual. No researcher was present while the home taping was being done. Once the children's oral production was transferred from a micro-cassette to a standard cassette, the language interaction was transcribed. The transcription was checked by at least one other Portuguese native-speaking researcher.

Twelve five-year-old native speakers (six of them on the Island of São Miguel, Azores, and six in Mainland Portugal) were also taped once for one and a half hours each. The native speaker sample was scored in the same way as the main sample (see below).

6. **Parents' Questionnaire.** The parents' questionnaire consisted of 34 items concerning the family background and language use of each child, the amount of exposure each child had had to printed materials in the home, and the parents' attitudes and expectations concerning the child's linguistic proficiency. Family background questions focused on place of birth, parental birthplace, education and occupation. Language use questions concerned the extent to which Portuguese and English were used with family members as well as with babysitters, nursery school and/or daycare personnel. The questions concerning the amount of exposure to printed materials asked parents if the child had been exposed to books or magazines in the home, if any of the parents read/told stories to the child, and the extent to which each child could read or recognize certain letters of the alphabet before going to school. The question concerning parents' attitudes and expectations asked parents to what extent they expected their children to become Portuguese-English bilinguals. The questionnaire was administered by two researchers at the end of the first visit (Year 1). One of the researchers asked the questions and filled in the results while the other took notes of all comments the parents made. The questionnaire was followed by an open-ended discussion that focused mainly on the parents' opinion of the heritage language program and language learning in general. This questionnaire was adapted from the Language Use Questionnaire previously developed for grade 7 students (see Chapter 7).

7. **Writing Vocabulary Test.** This test, also developed by Marie Clay, is described in Clay (1981). The test is individually administered to each child and consists of giving each child a piece of blank paper and asking him/her to write all the words he/she knows in five minutes, starting with the child's own name. Its purpose is to measure the amount of correctly spelled words each child can write, as a way of predicting his/her reading acquisition. In scoring, the following were taken into account: (1) all the words the student wrote, correctly or incorrectly spelled, including attempts at writing words that do not even exist in the lexicon (proper names were also counted); (2) the total number of words correctly spelled that are part of the lexicon. In the case of Portuguese it was decided that students should not be penalized for failure to use diacritics.

This test was administered in Portuguese to 26 grade 1 native speakers in the Azores and will be administered during Year 3 to the main sample of students in both Portuguese and English.

8. **Conversation Task.** This procedure was developed by project staff. Children were individually asked three questions designed to elicit particular verb tenses. All

answers from the students were recorded on tape, transcribed and scored. In Question 1, the tester asked the child for information about his/her family, such as where he/she lived, with whom, etc. This question aimed at eliciting from the child a series of obligatory contexts for the use of the present indicative. In Question 2 the tester asked the child about a trip the child had made by plane, boat or car. This question aimed at eliciting a small narrative of past events where a number of obligatory contexts requiring past tenses would occur. In Question 3 the tester asked the child what he/she would do, should the child win $3,000,000 in the lottery. This question aimed at eliciting a series of hypothetical statements where the following tenses would be required: conditional in English and conditional, imperfect subjunctive and/or imperfect with conditional or hypothetical aspect in Portuguese. The scoring consisted of a count of the number of obligatory contexts for tenses in each question in relation to the number of tense errors.

This task was administered in its Portuguese version to the 26 grade 1 native speakers in the Azores and will be administered during Year 3 to the main sample of students in both Portuguese and English.

9. **Gates-McGinitie Reading Comprehension Test.** This test, currently used in primary schools across Canada, will be administered to the main sample in Year 3.

3. SCORING PROCEDURES AND PRELIMINARY RESULTS FOR HOME INTERACTION DATA

Quantitative analyses of the data have not been carried out to date. The Year 1 home interaction data, however, have been transcribed and scored. The scoring scheme and preliminary trends are described below.

The spoken discourse of each child was transcribed, divided into communication units (C-units) and scored for grammar and pragmatics. In the following sections we will describe the criteria used for dividing spoken discourse into C-units, as well as the criteria used for the grammatical and the pragmatic scoring. Examples of the scoring procedures are taken from the profiles of the four middle-class Azorean children.

3:1 Criteria Used to Divide Spoken Discourse into C-Units and Words per C-Unit

For the purpose of our analysis we considered C-units as being constituted by an uninterrupted segment of speech that could not be further divided without loss of its essential meaning. Therefore, we disregarded in our C-unit count certain repetitions, fillers and edit mazes.[1] It should be noted, however, that in our sample there was no method of controlling the amount and the type of interaction between the target student and the other participants.

Walter Loban (1963) divides a discourse sample into C-units based on a series of answers to specific questions asked by an interviewer. This might account for his definition: "In actuality the communication unit in this research proves to be the grammatical independent clause with any of its modifiers" (Loban 1963, p.7). In our case, due to the natural environment in which spoken communication took place, we not only divided C-units according to Loban's method, i.e., a main clause plus the possible subordinate or coordinate clauses essential to its meaning, but had to deal with a series of interrupted utterances whose segments we had to consider as C-units, since they were delimited by two silences or two long pauses.[2]

If we consider the following example (where C-units are divided by slash marks):

A -- /John!/

B -- What do you want?

C -- /Come here !/

we could see that, had speaker A not been interrupted by speaker B, we would have one C-unit instead of two.

Following Loban, we also counted as C-units one-word utterances in cases where the isolated word could have been replaced by a main clause. This not only applies to questions capable of eliciting "yes" or "no" answers, but also to answers of the following type:

B -- How old are you?

A -- /Five./

B -- What's your name?

A -- /John./

After dividing the child's spoken discourse into C-units we counted the number of words per C-unit as an index of syntactic sophistication. Every time a simple preposition contracted with an article, a demonstrative pronoun or an adverb, e.g., "no" (in the), "daquele" (of that one over there), "daqui" (from here), we counted it as one word since it is generally perceived as forming a unit of meaning. However, prepositional locutions were counted according to the number of words that form them, e.g., "ao pé de" (close to) -- 3 words, "dentro de" (inside of) -- 2 words. Pronominalized direct or indirect objects hyphenated with the verb were counted as one word, e.g., "dei-o" (I gave it) -- 2 words, "dei-lhe" (I gave him) -- 2 words.

Table 1, p. 182, indicates the number of C-units and the average number of words per C-unit for each of the four middle-class urban children in the Azores. It is clear that considerable variation exists in the amount of language produced by the children.

3:2 Grammatical Scoring

To evaluate and describe children's acquisition of grammatical proficiency, we designed a scoring scheme that would assess both the nature and number of errors as well as correct uses: e.g., number of verb errors in relation to number of obligatory contexts for verbs. The following specific categories were scored:

1. Word order

2. Degree of adjectives
3. Omission, addition or substitution of any word
4. Lexicon (related to the use of verbs)
5. Prepositions and prepositional locutions
6. Verbs

Word order. When examining errors related to word order, it was noticed that they occurred only in constructions requiring the inversion of direct or indirect object pronouns as well as reflexive particles. Thus we counted: (1) the number of obligatory contexts requiring the type of constructions mentioned above and (2) the number of word order errors occurring in such constructions.

In Portuguese, pronominalized direct and indirect objects usually follow the verb, thus conforming to the most common type of word order in Romance languages, i.e., Subject + Verb + Direct/Indirect Object. This is the case for many affirmative clauses in Portuguese such as:

Eu	dei-lhe	o livro
S	V IO	DO
I	gave him	the book

Eu	dei-o	à Maria
S	V DO	IO
I	gave it	to Maria

However, there are several cases when pronominalized direct and indirect objects have to precede the verb:

(a) whenever a clause contains a negative word (not, never, anything, etc.), e.g.,

Eu	não	lhe dei	o livro
I	did not	give him	the book

Eu	nunca	lhe dei	o livro
I	never	gave him	the book

(b) whenever a clause starts with an interrogative or indefinite pronoun, e.g.,

Quem	lhe deu	o livro?

Who	gave him	the book?

Alguém	lhe deu	o livro
Someone	gave him	the book

(c) whenever a clause is introduced by a word or phrase expressing desire, e.g.,

Oxalá	ela	lhe dê	o livro
I wish	she	would give him the book	

(d) in subordinate clauses preceding or following a main clause, e.g.,

...quando	eu	lhe dei	o livro ...
...when	I	gave him	the book ...
...sempre que	eu	lhe dava	o livro ...
...whenever	I	gave her	the book ...
...se	eu	lhe der	o livro...
...if	I	give her	the book...

(e) whenever the verb is preceded by certain adverbs or adverbial locutions and both verb and adverb/adverbial locution belong to the same C-unit, e.g.,

Eu	já	lhe dei	o livro
I've	already	given him	the book

Note that the same grammatical rules apply to reflexive pronouns, e.g.,

Ele	esqueceu-se
He	forgot (himself)

Ela	não	se esqueceu
She	did not	forget

Due to the complexity of rules concerning the placement of pronominalized direct and indirect objects as well as reflexive particles, some of the five-year-old native speakers (in the Azores) made some mistakes regarding word order. The following are examples taken from the Azorean transcripts:

Child 1

"Eles já foram-se embora" instead of "Eles já se foram embora"

(They already went out)

"o papá não vai-te dar..." instead of "o papá não te vai dar"

(papa is not going to hit you)

Child 4

"as minhas bonecas também pode-se" instead of "as minhas bonecas também se podem"

(my dolls can also...)

According to Clark (no date) the placement of pronominalized direct and indirect objects in Romance languages is a product of later acquisition: "fairly late is full mastery of the different word orders required with direct and indirect object clitic pronouns, in affirmative versus negative and in indicative versus imperative sentence forms" (p. 14).

Degree of adjectives. When examining errors related to degree of adjectives, we noticed that they only occurred in the Azorean sample when the children used irregular adjectives such as: "bom" (good), "melhor" (better), "mau" (bad), "pior" (worse). Thus we counted all the occurrences of irregular adjectives as well as number of mistakes concerning them. The following is an example taken from the Azorean transcripts:

Child 1

"ainda mais melhores" (even more better)

It should be noted that the simple superlative was never used by any of the four middle-class Azorean children. Due to the fact that the forms of the simple superlative in Portuguese are very close to Latin, e.g., "péssimo" (very bad), "felicíssimo" (very happy), "paupérrimo" (very poor), such forms appear to be a product of relatively late acquisition. The forms of the compound superlative are far more current in spoken language, e.g., "muito mau" (very bad), "muito feliz" (very happy), "muito pobre" (very poor).

Omission, addition or substitution of any word. In this category we listed occurrences each time they did not fall within the other categories of our scoring scheme. Thus, omissions or additions of verbs or prepositions were counted as errors respectively under Verbs and Prepositions. It should be noted, however, that the majority of such errors are related to the use of pronominalized direct and indirect object pronouns.

Lexicon. Even though errors related to lexicon would not normally be counted under grammar, we decided to count lexical errors related to the use of certain verbs in our grammatical scoring. These errors are in fact more of a semantic nature but nevertheless they make the sentences where they occur look somewhat ungrammatical. Child 3 has, in certain cases, difficulty using the verb pair "dizer/perguntar" and its alternative "pedir" (to tell/to ask). The following are three examples taken from his transcript, (a) and (c) presenting incorrect usages and (b) a correct usage of the same pair.

(a) "Vou dizer à mamã que horas são. Mae, que horas são?"

(I am going to tell mama what time it is. Mother what time is it?)

(b) "Eu vou perguntar uma coisa ao Miguel. Que horas são?"

(I am going to ask Miguel something. What time is it?)

(c) "Eu digo à mamã... vou pedir à mamã que não quero mais."

(I'll tell mama... I am going to ask mother that I don't want any more)

Note that ask/tell do not function the same way as in English. While in English ask can be used: (1) to obtain information, (2) to command politely or request; in Portuguese, "perguntar" (to ask) means only to obtain information. To command politely, one would have to use another verb: "pedir". "Dizer" (to tell) means, as in English, (1) to give information and (2) to give a command.

Child 1 also had difficulty with another pair "saber/conhecer" both translated in English by the verb to know. The following is an example taken from Child 1 of the incorrect usage of the verb "conhecer":

"...não estás conhecendo que isto está a gravar?"

(... don't you know this is taping?)

The Portuguese usage of "saber/conhecer" is similar to the French usage of "savoir/connaître". "Saber" means generally to have knowledge gained through learning, memorization or studying, while "conhecer" is used generally with a sense of being familiar with, through having seen or met. Thus "conhecer" is always used when referring to people.

Prepositions and prepositional locutions. In this category we counted all the obligatory contexts for their usage, including and excluding the preposition "a". We decided to count obligatory contexts excluding "a" due to the fact that Azorean speakers, as opposed to those in Mainland Portugal, tend to form the progressive aspect in verbal conjugations in a different way. While in the Mainland the progressive aspect is formed by using a form "estar" (to be) followed by the preposition "a" plus an infinitive, the Azoreans tend to use a form of "estar" plus the gerund. Both forms are recognized by Portuguese grammarians as belonging to "standard Portuguese". After having counted all the obligatory contexts for each child, we counted the number of errors.

To define what constitutes a preposition and a prepositional locution we followed the definitions of Cunha and Cintra (1984) in Nova Gramática do Português Contemporâneo. Prepositions are defined as "invariable words that relate two terms in a clause in such a way that the sense of the first (antecedent) is explained or completed by the second (consequent)".[3] Prepositional locutions are defined in the same grammar as "being constituted by two or more lexical items, the last one being a preposition".[4] This grammar, as well as other Portuguese grammars, gives a fairly extensive list of prepositions and prepositional locutions and explains their usage. It should be noted, however, that "mais" is not listed as a preposition even though it is used in colloquial expressions of the type: "Eu vou mais ela" (I am going with her), i.e., in expressions where "mais" is a synonym of "com" (with).

The native Portuguese-speaking children in the Azores did not have too many problems using prepositions and prepositional locutions. They were, in fact, able to transmit through them correct notions of space, time and relation. However, Child 4 seemed to have some difficulties with prepositional locutions. The following are examples taken from the transcript:

"vou passar-me atrás da mesa" instead of "vou passar-me por detrás da mesa"

(I am going to go through behind the table)

"a avó vem para atrás de mim" instead of "a avó vem para trás de mim"

(grandmother will be behind me)

For each child we counted the range of prepositions and prepositional locutions that were correctly used. As with the obligatory contexts we gave a number including and excluding "a".

Verbs. Under verbs we counted all the obligatory contexts for finite verbs, i.e., we did not count infinitives or isolated gerunds or past participles. Use of a tense such as the present progressive, the immediate future or any tense conjugated with an auxiliary was counted as one instance, e.g., "estou comendo" (I am eating) counted as one form of the verb to eat; "ele vai ver" (he is going to see) counted as one form of the verb to see. Afterwards, we counted the number of errors: wrong person and/or agreement and wrong tense within the sentence.

In order to assess the extent to which five-year-old native speakers use regular verbs belonging to all three conjugations as well as irregular verbs, we counted the number of regular and irregular verbs used.

In Portuguese the first conjugation ends in -ar, e.g., "falar" (to speak), "cantar" (to sing); the second in -er, e.g., "comer" (to eat), "bater" (to hit) and the third one in -ir, e.g., "fugir" (to run away), "dormir" (to sleep). The first conjugation was by far the most used due to the fact that in Portuguese the majority of regular verbs end in -ar. We also counted the number of irregular verbs including and excluding "ser/estar" (to be) since the latter are the ones more commonly used. We based our criteria for defining an irregular verb on Cunha and Cintra (1984). Our data show a large number of irregular verbs being used by the children because, as in many other languages, some of the most commonly used verbs are irregular. The following lists the lexical range of irregular verbs used by all four middle-class children in the Azores:

cair	(to fall)	poder	(can)
dar	(to give)	pôr	(to put)
dizer	(to say)	querer	(to want)
estar	(to be)	saber	(to know)
fazer	(to do)	sair	(to go out)
haver	(there be)	ter	(to have)
ir	(to go)	ver	(to see)
ouvir	(to hear)	vir	(to come)

As far as range of tenses was concerned, we listed all conjugated forms disregarding, in this case, such labels as 'mode' or 'aspect'. Thus, the imperative and the subjunctives were counted as tenses. It should be noted that in Portuguese all the subjunctives -- present, imperfect and future -- are used largely independently of the level of formality. The conditional and the simple future did not occur in the Azorean sample, possible due to the fact that these tenses are mostly used in formal oral and written language and can be replaced, respectively, by the imperfect and by the immediate future.

It is interesting to point out that some children used a verb tense -- personal infinitive -- that is used mainly for stylistic purposes. This tense, found only in Portuguese, functions as a way of personalizing certain infinitives, especially in clauses where the subject is omitted, e.g.,

"isto é para veres"

(this is for you to see)

The contexts that occurred in the small Azorean sample did not require the use of compound tenses.[5] However, the fact that the children could use the subjunctive mode as well as the personal infinitive might lead us to infer that the major tenses in the Portuguese verbal system have been mastered. Only one error concerning the use of the subjunctive was encountered in the four Azorean transcripts:

Child 1

"antes que ele vem (present indicative)" instead of "antes que ele venha (present subjunctive)"

(before he comes)

The other verb errors were mainly related to the use of wrong person, e.g.,

Child 4

"está aqui os raiozinhos"

literally, (the little rays of sunshine is here)

Sometimes Child 1 would use a formal register (i.e., 3rd person singular) to address a younger brother. We did not consider this as a grammatical error, nor should it be considered, in this case, as a sociolinguistic error. If one examines the contexts where this occurs one can notice that switch of registers, i.e., informal to formal, is used here to express assertiveness or anger, thus imitating the change of registers used by adults when speaking to a very young child. Frequently, adults would address a child as if the child were a 3rd person, e.g., "O Joao esteja (3rd person singular) quieto." (John stay still!) The following are two examples of this usage:

Child 1

"Honório! Honório! Não foge! (3rd person singular)"

(Honório! Honório! Don't run away!)

"Honório! Não vai (3rd person singular) muito para aí."

(Honório! Don't go over there.)

Few dialectal traits were found in the Azorean sample. Nevertheless, we found some occurrences of the use of TER (to have) for HAVER (there be), e.g.,

Child 3

"tem lá três rapazes"

literally, (has three boys there)

It should be noted, however, that such usage is accepted in spoken Brazilian Portuguese. As in scoring schemes developed for other Development of Bilingual Proficiency studies, we did not count dialectal variants as errors.

3:3 Pragmatic Scoring

Based on criteria elaborated by Damico, Oller and Storey (1983), we developed a pragmatic scoring scheme applicable to home interaction data. According to Damico et al, the main emphasis of a pragmatic assessment consists in analysing how efficiently meaning is conveyed between speaker and audience, this including both the use of language in its social context and also the ability to map specific utterances onto particular contexts of experience.

In view of the fact that our sample was not recorded in an interview situation, we disregarded three categories from the Damico et al. scoring scheme: (a) delays before responding, (b) poor topic maintenance and (c) need for repetition on the part of the interviewer. We also blended two of the Damico et al. categories into one; i.e., where the authors have two categories, (a) linguistic nonfluencies and (b) revisions,[6] we decided on a single category, Nonfluencies, subdivided in the following way:

1. Nonfluencies
    - (a) audible/silent pauses
    - (b) inappropriate use of repetitions
    - (c) filler words and phrases

(d) edit mazes

The other categories in our scoring scheme are:

2. Inappropriate or off-topic responses
3. Nonspecific vocabulary
4. Inappropriate use of anaphora and/or cataphora

For each one of the above categories we counted the number of occurrences. This was done directly from the tape while reading the transcripts.

The specific categories and subcategories used in our scoring scheme are as follows:

Nonfluencies. In this category we scored for:

(a) Number of audible or silent pauses in the middle of utterances or a C-unit or in between syllables belonging to the same word. By audible pauses we mean hesitations of the type: "uh", "mm", and the like. It should be noted, however, that native speakers frequently hesitate in the middle of an utterance or of a C-unit, or in the middle of a word. Nevertheless, such hesitations or pauses are, in the case of a native speaker, virtually always accompanied by an appropriate tone of voice able to convey additional meaning such as: physical or mental fatigue, sudden loss of memory, etc.;

(b) Number of inappropriate uses of repetitions. This category was used each time the child's tone of voice indicated that the repetition was indicative of difficulty in conveying a certain linguistic message;

(c) Number of non-native-like uses of filler words and phrases. By these we mean the incorrect usage of certain phatic words or phrases at the beginning or at the end of a prosodic unit, the main function of which is not only to maintain contact between sender and receiver, but also to emphasize or assert certain statements. Speakers of a second language or of a first language in a bilingual context, when using such filler words or phrases, tend either to code-switch or to transfer them from one language to the other by means of a literal translation;

(d) Number of edit mazes. This metaphor used by Loban (1963) defines mainly a series of false starts, as well as incoherent and noncohesive segments. He defines edit mazes as a "series of words or initial parts of words which do not add up either to meaningful communication or to structural units of communication... they are unattached fragments or a series of unattached fragments which do not constitute a communication unit and are not necessary to the communication unit" (Loban 1963, p. 8). In our scoring scheme we counted as edit mazes any false starts constituted by two or more words or any discourse segment constituted by a series of words or word fragments unable to convey meaning. Note that once a given segment was categorized as an edit maze, it was not counted under number of audible/silent pauses, inappropriate use of repetitions or filler words or phrases.

Inappropriate or off-topic responses. In this category we counted all instances where such responses followed direct or indirect questions, e.g.,

A Where are you going?

B (1) Did you see her over there?
(2) I saw John yesterday.

Note that native speakers can also, like speaker B above, use inappropriate or off-topic responses as a strategic way of avoiding or switching topics. Thus, the scorer has to take into consideration a wider context to determine if such off-topic responses are, in fact, inappropriate.

Nonspecific vocabulary. We scored in this category only the instances where children used generic terms such as "thing" or "stuff" in lieu of common vocabulary items. While Damico et al. include under this heading the inappropriate use of cataphora (i.e., deictic expressions when no antecedents have been provided by the speaker and when the listener has no way of knowing what is being referenced), we decided to include inappropriate use of cataphora under a more specific category.

Inappropriate use of anaphora and cataphora. Under this heading we counted all the instances where: (1) subsequent references refer inappropriately to previous information (wrong use of anaphora), and/or (2) when no antecedents have been provided and the listener cannot understand what the speaker is referring to (inappropriate use of cataphora).

3:4 Pragmatic Performance of the Azorean Sample

Even though some children spoke more than others, they all interacted with several people. The people with whom each of the four Azorean children interacted during the taping session are listed below:

Child 1: father
mother
uncle
grandmother
younger brother
female adult
other children

Child 2: father
mother
sister
female adult

Child 3: mother
sister

brother
male adult
other children

Child 4: father
mother
grandmother
brother

These children, not surprisingly for native speakers, scored quite well in pragmatics. In fact, apart from some unnecessary repetitions, one inappropriate or off-topic response in the whole sample and some edit mazes, they encountered no other pragmatic problems. The transcript of Child 1 shows the occurrence of seven edit mazes. The following are two examples of these:

"Eh, Nuno, o que é que estás a... porque é que ainda estás aí a chorar?"

(Hey Nuno, what are you... why are you still crying?

"O Miguel andava sempre com aquilo na mão e o papá disse assim ... uh...ele fez a...uhm e o Miguel brigou com ele."

(Miguel was always carrying that [a paint gun] in his hand and Dad said that ...uh...he did the...uhm and Miguel fought with him.)

4. HOME INTERACTION OF 4-YEAR OLD PORTUGUESE BACKGROUND CHILDREN IN TORONTO: PRELIMINARY TRENDS

All the children in the main sample (Toronto) were taped for a one and a half hour period in the home. They were encouraged to speak whatever language they wanted (English or Portuguese) and they were also told to go about their normal routines as usual. While transcribing the tapes, it became obvious that some children were being coaxed by family members into speaking Portuguese. In some cases, the interaction between parents and children was carried out in both languages; for example, children might be addressed in Portuguese and they would systematically answer in English.

Whenever English was the language mostly used, it tended to be used almost exclusively. In the case of Portuguese, the results were more diverse; in some cases, Portuguese was used more often only by a small margin: e.g., 50.23%, 51.56% and 52.11%. The range for the percentage of English was between 99.01% and 20.41%. In Portuguese the range was between 79.58% and 0.67%.

4:1 Grammatical Performance

Syntactic sophistication. With respect to the index of syntactic sophistication, i.e., average words per C-unit, we considered only the averages of children that produced more than 15 C-units, be it in Portuguese, English or mixed code. Among the native speakers from the Azores (whose age is on average one year older than the Year 1 Toronto sample), the averages for syntactic sophistication range from 3.90 to 5.05, the majority being in the 4. range. In the main sample in Toronto, greater variation was found. The averages for syntactic sophistication range, in Portuguese, from 1.44 to 4.85, while in English the averages range from 1.16 to 5.74. However, the great majority of children showed a percentage in the 3. range. Note that in the case of English, there are no native speaker data for comparison purposes. Concerning the use of mixed code (i.e., C-units where both Portuguese and English occurred), only four children within the main sample produced more than 15 C-units. Here the averages were considerably higher, ranging from 5.53 to 6.80. In certain cases, the measure for syntactic sophistication was considerably higher in mixed code, probably due to the fact that the students did not produce too many C-units of this type. Thus, for a student that only produced six C-units in mixed code, we have a measure of syntactic sophistication of 10.00.

Word order. With respect to word order errors, it was noticed that among the native speaker (Azorean) sample, they occurred only in constructions requiring the inversion of direct and indirect object pronouns as well as reflexive particles. Among the main sample it was noticed that even though these children showed other types of word order errors, the vast majority made errors each time the direct or indirect object pronouns, as well as reflexive particles, were required. As far as English was concerned, only a few children in the main sample made errors regarding word order. It should be noted that in Portuguese, the rules that govern the placement of pronominalized direct or indirect object pronouns are quite complex.

Degree of adjectives. When examining all errors related to degree of adjectives we noticed that, among native speakers, they only occurred when children used irregular adjectives such as: "bom" (good), "melhor" (better), "mau" (bad) and "pior" (worse). Among the children in the main sample, this type of error seldom occurred either in English or Portuguese, due to the fact that there were practically no obligatory contexts requiring the use of these forms.

Omissions, additions or substitutions. Native speakers only made errors related to the use of pronominalized direct and indirect object pronouns. The students in the main sample showed the same pattern, although errors were more frequent. Omission of conjunctions was also more freqent in the main sample in the case of Portuguese. As far as English is concerned, there were no major problems in this category.

Lexicon. Although difficulties with the pair "dizer/perguntar" (ask/tell) did not appear in the main sample (Toronto), these children showed the same problems with "saber/conhecer" (to know) as have already been noted in the Azorean sample. In addition, unlike the native speakers, the students in the main sample had problems with the use of "ser/estar" (to be) and with "estar/ir" (to be/to go) when used as auxiliaries. In English usage some errors did occur concerning, in particular, the following pairs: "learn/teach" and "win/beat".

Prepositions and prepositional locutions. The students in the main sample, like those in the Azores, showed no significant problems with the use of prepositions and/or prepositional locutions in Portuguese or in English. Only eight students, out of a sample of 20, made mistakes pertaining to this category, either in English or in Portuguese.

Verbs. Under this category it was noted that the native speaker children made very few errors. In fact, in the whole native speaker sample, only one child made six 'mistakes' concerning verbs and of these six 'mistakes', four could be considered as

Azorean dialectal variants. As far as the range of tenses is concerned, the native speakers used quite a variety of tenses. It is interesting to point out, for example, that several children used the personal infinitive, a Portuguese verb tense that is mainly used for stylistic purposes. The subjunctive was also used correctly.

In the main sample, the children had more problems with verbs, especially in Portuguese. There were three children with a high number of errors in this category (15, 17 and 15); however, it should be noted that these children produced more than 300 C-units each. In general the children made relatively few verb errors. In the whole sample (20 children), six children made no errors and another six had only one error each. In English, the children in the main sample made a few errors regarding the use of verbs. There were two students with 14 errors each; however these two children had produced considerably more than 300 C-units. There were six children with no errors and another six with one error each. As far as range of tenses in the main sample was concerned, the children used more or less the same range of tenses in Portuguese as the native speakers. Even though the tenses that occurred more often were the present indicative, the simple past and the imperative, the tenses of the subjunctive as well as the personal infinitive also occurred. Note that as with the native speakers, the simple future and the conditional never occurred in Portuguese. In English the children used a large variety of tenses, including a variety of modalities as well as the conditional and the simple future; they also did not show any particular problems with the use of compound tenses; they used present progressive, present perfect and past progressive several times. It is clear that the children in the main sample have acquired a large variety of verb tenses in both English and Portuguese.

Interference. Under interference, we decided to consider all types of errors that were either of a simple lexical nature or transpositions of certain grammatical constructions from English to Portuguese or vice versa. It is interesting to note that the students in the main sample did not have too many mistakes related to interference; in the whole sample, no examples of Portuguese in English were found and only three students showed one or two mistakes concerning interference of English in Portuguese. As far as lexicon was concerned, only one mistake of this nature was found: "mechim" (machine) instead of "máquina". Grammatical interference in Portuguese has to do with the use of certain English constructions of the type "comes back" used in Portuguese as "vem para trás" instead of "volta." Note that this feature is also found in French among French immersion students and is quite common among Franco-Ontarian speakers.

4.2 Pragmatic Performance

As noted above, the native speaker children performed well with respect to pragmatics. They showed only a few mistakes related to inappropriate use of repetitions, inappropriate or off-topic responses and edit mazes. Among the children in the main sample, there were more problems concerning pragmatics, even though the number of errors was fairly low. If we look at the number of errors for the whole sample, we could say that they were virtually absent in mixed code utterances. There seem to be more errors in English than in Portuguese; also the diversity of errors is greater in English. Even though the number of errors in both English and Portuguese is higher in edit mazes, there are other areas in English where there is an incidence of problems that do not occur in Portuguese, such as the use of non-native-like filler words or phrases and non-specific use of vocabulary. Note, however, that in the whole sample the children used more English than Portuguese (59% versus 41%); this might account for the fact that there were more errors in English than in Portuguese.

4.3 Conclusion

In conclusion, the language skills of the minority children in the main sample appear to be following a similar developmental pattern as is the case for the Azorean children whose language is developing in a majority context. When the Year 2 (five-year-old) transcripts are analysed, the longitudinal trends and Toronto-Azorean comparisons can be described in more detail than is presently the case.

Footnotes

1. Even though the number of unnecessary repetitions and edit mazes shows in the pragmatic scoring scheme, the number of fillers does not, since we only counted the number of non-native-like uses of filler words and phrases. However, correctly used filler words or phrases were disregarded when we counted the number of words per C-unit.

2. These are the 'natural pauses' a native speaker would use in oral discourse.

3. Our translation.

4. Ibid.

5. The use of compound tenses would occur if we had had a number of contexts requiring the narration of past events.

6. By 'linguistic nonfluencies' Damico et al. refer to the disruption of speech production due to a large number of repetitions. By 'revisions' they mean the instances when speech production is broken up by numerous false starts or self-interruptions.

References

Clark, Eve. V. **The Acquisition of Romance with Special Reference to French.** Stanford, n.d.

Clay, M.M. **The Early Detection of Reading Difficulties: A Diagnostic Survey.** 2nd ed. London: Heinemann, 1981.

Clay, M.M., Gill, W.M., Glynn, E.L., McNaughton, A.H. and Salmon, K.W. **Record of Oral Language.** Wellington: New Zealand Educational Institute, 1976.

Cunha, C. and Cintra, L. **Nova Gramatica do Português Comtemporâneo.** Lisbon, 1984.

Damico, J., Oller, J.W. Jr. and Storey, M.E. The diagnosis of language disorders in bilingual children: Pragmatic and surface-oriented criteria. **Journal of Speech and Hearing Disorders,** 1983, **48,** 385-94.

Harris, D.B. **Children's Drawings as Measures of Intellectual Maturity: A Revision and Extension of the Goodenough Draw-a-Man Test.** New York: Harcourt, Brace & World, 1963.

Loban, W. **The Language of Elementary School Children.** Champaign, Illinois, 1966.

Loban, W. **Problems in Oral English: Kindergarten to Grade Nine.** Champaign, Illinois, 1966.

Nunes, Carmen, et al. **Nova Gramatica de Português.** Lisbon, 1981.

Table 1

Number of C-units and Average Number of Words per C-unit for Four Azorean Children

Child	No. of C-units	Average no. of words per C-unit
1	563	4.442
2	23	4.260
3	123	4.268
4	177	4.022

Chapter 9

AGE OF ARRIVAL, LENGTH OF RESIDENCE, AND INTERDEPENDENCE OF LITERACY SKILLS AMONG JAPANESE IMMIGRANT STUDENTS

Jim Cummins and Kazuko Nakajima

1. BACKGROUND

In this study we explore several issues related to the development of bilingualism among Japanese children in Canada who are being educated under relatively favourable conditions. The families have good economic circumstances, English-as-a-second-language support is provided in most schools, and the development of Japanese academic skills is promoted by the Japanese school on week-ends. The study investigates the influence of age on arrival and length of residence in Canada on the development and maintenance of children's Japanese and English academic proficiency. The relationship between their Japanese and English academic proficiency is also investigated. Specifically, we asked the following questions:

- To what extent and how rapidly does children's L1 academic proficiency decline over time?

- What are the prospects for continued L1 development among children who arrive in Canada at different ages (e.g. 4 years as compared to 8 years)?

- How long does it take Japanese children to acquire age-appropriate English academic skills and are there differences in the rate at which these skills are acquired by older and younger students?

- What is the relationship between children's L1 reading and writing skills, on the one hand, and their acquisition of English reading and writing skills on the other?

1:1 The Policy Context of the Study

A greater understanding of these issues has clear importance for policy-makers in both the home and host countries (i.e. Japan and Canada in this case). For example, the difficulties faced by returnee Japanese children attempting to reintegrate into Japanese schooling and society have become increasingly evident and similar issues are being debated in many other countries (e.g. Greece, Portugal) that are experiencing the return of guest-workers' children as economic conditions deteriorate in the host countries. If we understood what factors helped children maintain their L1 proficiency while they were acquiring the language of the host society, we could better intervene to promote the development of full bilingualism. Not only would the personal opportunities available to the individual student increase if full rather than limited bilingualism were developed, but also from an economic point of view, these fully bilingual children would constitute a major resource for future trade and international relations.

1:2 The Theoretical Context

Many of the issues we investigate in this study have long been debated by theorists of L2 acquisition around the world. For example, it has frequently been assumed that younger learners had a significant advantage over older learners in acquiring L2 proficiency; theorists argued that there was a pre-pubertal "optimal age" or "critical period" for acquiring L2 proficiency and native-like skills could not be acquired after that time. More recent studies, however, cast doubt on the notion of a critical period, at least for most aspects of L2 proficiency (for example, see Cummins 1981; Harley 1986 for reviews).

A related issue concerns the influence of L1 proficiency on the development of L2 skills. In many countries, students from minority language backgrounds were encouraged (often through physical punishment) to replace their L1 by the language of the school. This practice was rationalized by educators on the grounds that maintenance of L1 would interfere with children's acquisition of English and reduce their chances of academic success. These assumptions are still very much in evidence among educators of minority students; for example, in discussing with parents the results of a previous study we conducted with Japanese students in Toronto (Cummins et al 1984), several parents (whose English was minimal) told of Canadian teachers advising them to speak only English with their child.

In contrast to these assumptions, it has been proposed (e.g. Cummins and Swain 1986) that there is a strong positive relationship between the development of L1 conceptual language skills and the subsequent acquisition of L2 conceptual skills. In other words, L1 conceptual proficiency provides the foundation for the development of L2, or expressed differently, L1 and L2 skills are interdependent in that they are manifestations of a common underlying proficiency. Support for this theoretical proposition comes from several sources; for example, research on bilingual education, correlational research on the relationship between L1 and L2, research showing advantages in acquiring L2 academic skills for older L2 learners whose L1 proficiency is better developed than that of younger learners (see Cummins and Swain 1986 for a review).

The theoretical notion of a common underlying proficiency among bilinguals is related to distinctions that have been proposed regarding the nature of language proficiency (Cummins and Swain 1986). Specifically, the importance from both practical and theoretical perspectives of distinguishing conversational from academic language skills has been emphasized. This distinction has been discussed in terms of the range of contextual cues that support the meaning (Cummins 1984). Specifically, conversational skills are described as "context-embedded" in that the communication is embedded in a context of linguistic, paralinguistic, and situational cues (e.g. intonation, gestures, pointing, etc.) that support the meaning. Academic or literacy-related language skills, on the other hand, are described as context-reduced in that the range of cues to meaning is considerably reduced in comparison to face-to-face conversation. For example, in reading a difficult text, the individual must rely almost exclusively on linguistic cues (e.g. syntax, semantics) to discover the meaning.

The relevance of the context-embedded/context-reduced distinction can be seen in the fact that it takes immigrant students considerably longer (five to seven years on average) to attain age-appropriate levels in L2 academic skills than in L2 conversational skills (approximately two years on average) (Cummins 1981). If educators assume that all difficulties with L2 (e.g. English) have been overcome when the minority child appears to converse and understand English well, serious errors can be made in interpreting children's English academic performance in the classroom or on tests.

The present study further investigates both the interdependence hypothesis and the relationship between age and second language acquisition, with a focus on the development of reading and writing skills. Previous studies investigating these issues, with few exceptions, have not considered the cross-lingual relationships among L1 and L2 writing skills. An investigation of the relationships between Japanese and English proficiency appears to provide a stringent test of the interdependence hypothesis since the two languages have little in common at a surface structure level.

2. METHODOLOGY

Subjects in the study consisted of 273 students between grades 2 and 8 attending the Japanese School of Toronto Shokokai Inc. Students were tested in May and June 1984 with measures of reading and writing in both Japanese and English. The reading comprehension subtest appropriate to students' grade level of the Gates-McGinitie Reading Test was given to all students who had been in Canada for at least six months as a measure of English reading skills. The Kyoken Standardized Diagnostic Test of Reading Comprehension published by the Research Institute for Applied Education in 1981 was given as the measure of Japanese reading skills. In addition, a letter writing task in English and Japanese was administered to all children (see Appendix A, pp. 200-202).

Scores on the English and Japanese reading tests were converted to T-scores to permit comparability across grades with the influence of age removed. In addition, English grade equivalent scores were used in some analyses as an approximate index of students' absolute level of English reading skills. A variety of indices of writing skills in Japanese and English were assessed; specifically, the following indices were assessed for English:

Words in T-units
Number of T-units
Words in mazes
Number of mazes
Number of spelling errors
Number of other grammatical errors

All of the above indices were assessed by simple counting of occurrences. In addition, several dimensions of students' writing were assessed by means of global rating scales. The scale for the first three variables ranged from 1 (inadequate) through 5 (excellent) while for the final variable the response of the reader was rated as low, medium, or high. The following indices were assessed:

Quality and range of content
Organization and presentation of content
Style and appropriateness
Interest of letter to rater

Twenty-seven Japanese variables were scored. Eighteen of these involved ratings related to the quality of students' writing while the remaining nine involved a checklist of specific errors that students might make. The variables are as follows:

1. Total number of letters used

2. Total number of errors in the use of Hiragana letters

3. Total number of Chinese characters used

4. Total number of errors in the use of Chinese characters

5. Spelling (Katakana)

   (0) no use or inappropriate use of Katakana letters
   (1) some unconventional use of Katakana letters
   (2) proper use of Katakana letters
   (9) unable to judge

6. Correctness in morphological structure (particles, auxiliary verbs)

   (0) many errors
   (1) occasional errors
   (2) proper use
   (9) unable to judge

7. Level of syntactic structure sophistication

   (0) generally only simple structures used
   (1) few complex structures attempted
   (2) many complex and long sentences used
   (9) unable to judge

8. Punctuation, spacing and handwriting

   (0) carelessness in punctuation, spacing and handwriting
   (1) occasional carelessness in punctuation, spacing and handwriting
   (2) proper use of punctuation, neatly spaced and easy to read

9. Organization of ideas and paragraphing

   (0) ideas are poorly organized and no attempt at paragraphing
   (1) some attempt at organizing ideas and paragraphing
   (2) ideas are well organized and paragraphs are clearly formed
   (9) unable to judge

10. Structural and lexical cohesion (use of connectors and anaphora)

    (0) poor use of connectors and anaphora
    (1) some attempt at using connectors and anaphora
    (2) appropriate use of connectors and anaphora

(9) unable to judge

11. Richness in content (vocabulary and information communicated)

(0) vocabulary selection is limited and the content is poor
(1) the content is average, with adequate vocabulary selection
(2) rich in content with the use of specific vocabulary
(9) unable to judge

12. Use of expressions appropriate for letter-writing

(0) no attempt at using expressions appropriate for letter-writing
(1) some attempt at using expressions appropriate for letter-writing
(2) proper use of expressions appropriate for letter-writing

13. Use of letter-writing formulae (date, signature, title of the recipient)

(0) no attempt at using letter-writing formulae
(1) some attempt at using letter-writing formulae
(2) proper use of letter-writing formulae

14. Appropriate choice of topic for letter-writing (awareness of the recipient of the letter)

(0) poor choice of topic, poor sense of recipient
(1) the choice of topic is ordinary, with some sense of recipient
(2) topic is well chosen, with good sense of recipient
(9) unable to judge

15. Awareness of honorifics and polite expressions

(0) no awareness
(1) somewhat aware of honorifics and polite expressions
(2) clearly aware of honorifics and polite expressions
(9) unable to judge

16. Ease and confidence in writing

(0) extremely laboured in writing
(1) somewhat laboured in writing
(2) clearly at ease in writing
(9) unable to judge

17. Attitude towards the lack of language skills

(0) withdrawn and unwilling to write
(1) willing to write but only as much as they can write without mistakes
(2) willing to write even though there may be some mistakes
(9) not applicable

18. Influence of English (phonetically, orthographically, syntactically)

(0) traces of the influence of English
(1) no traces of the influence of English
(9) unable to judge

Checklist of Errors

19. Forget to put small "tsu"

20. Forget to put particles

21. Put an extra space

22. Forget to use titles (Mr., Miss, etc.)

23. Poorly organized passages

24. Use English words

25. Overuse "I"

26. Use English period (.), instead of Japanese sentence end symbol

27. Inappropriate choice of words

3. RESULTS

The means and standard deviations for length of residence (LOR), age of arrival (AOA), English T-score (ETS), English grade equivalent (EGE) and Japanese T-score (JAT) are presented in Table 1, p. 194, for the 226 students for whom complete data were available for the English reading and writing variables. It can be seen that the sample as a whole is performing close to Japanese norms with respect to Japanese reading skills, a finding which is consistent with data from similar students in the Cummins et al (1984) study. On average, students have been in Canada almost four years, but the variation is considerable, as indicated by the standard deviation of 39.5 months for LOR. The average age of arrival is six and one-half years. Students, on average, are performing at the 35th percentile (English T-score 41.4) in English reading but the variation here is considerably greater than for Japanese reading (standard deviation for T-scores of 16.8 v. 10.4).

The relationship between age of arrival and Japanese T-score, on the one hand, and English reading and writing indices, on the other, was investigated by means of various analyses. Table 2, p. 195, presents the zero-order partial correlations and partial correlations controlling for length of residence between the English variables, age of arrival (AOA) and Japanese T-score (JAT). The zero-order correlations clearly show the strong relationship between length of residence (LOR) and performance on the English reading measures. Indices of English writing performance (particularly quality, organization and proportion of spelling and other errors) are related to LOR to a lesser extent. Japanese reading performance (JAT) is negatively related to LOR indicating a decline in performance in relation to Japanese grade norms the longer the student is away from Japan. The negative correlation between JAT and LOR is less than the positive relation between English T-score (ETS) and LOR (-.40 v. .53) suggesting that Japanese reading skills decline at a slower rate than English reading skills are acquired. This is consistent with the findings of the previous Cummins et al (1984) study and is presumably due, in part at least, to the influence of the Japanese supplementary school in promoting Japanese academic skills. Japanese reading scores are positively related to age of arrival indicating that students who arrive at older ages tend to maintain Japanese reading skills better than those who arrive at younger ages. The relationship between AOA and JAT remains significant ($p < .05$), although considerably reduced, when LOR is controlled (partial $r = .14$).

Cross-tabulations of English T-score and English grade equivalent by LOR and AOA are shown in Tables 3 and 4, pp. 196 and 197. Figures 1 and 2, p. 199, show the same data graphically with the exception that cells for which the N is less than 10 have not been graphed. LOR has been broken into three categories (0-35 months, 36-48 months, and 49+ months) while AOA is divided into five categories (0-5 years, 6-7 years, 8-9 years, 10-11 years, and 12+ years). English T-scores increase with increases of LOR, as would be expected. This effect is independent of AOA as can be seen in Figure 1 where for AOA 6-7 and AOA 8-9 the effects of LOR are shown within AOA categories. The relationship between AOA and English T-score is less consistent than for LOR and English T-score. It is interesting to note that students who have been in Canada for four years are performing at grade norms in English reading. This figure is less than reported in Cummins (1981) for a sample of 1210 immigrant children from various backgrounds, but consistent with the data reported in Cummins et al (1984) for a similar sample of middle- to upper-class Japanese children.

In general, students who arrive before age 8 (i.e. those in categories 1 and 2) tend to perform better in English reading in relation to their respective grade norms; however, only for those who have been in Canada less than three years is there an evident decline in English T-score with increasing age of arrival (at least for AOA categories 1-4).

With respect to English grade equivalent, the effects of LOR are clearly shown in Table 4 and Figure 2. Age of arrival also has a clear impact within LOR categories 36-48 months and 49+ months. Children who arrive at older ages perform better, in absolute terms, than those who arrive at younger ages with the same length of residence. However, for LOR category 0-35 months, the trend for older children to perform better in English reading is weaker, the only obvious difference being between AOA categories 10-11 and 12+. This may be due to possible confounding between LOR and AOA within the LOR 0-35 months category; at this early stage of acquisition small differences in length of residence between AOA categories could significantly skew the pattern of results.

Regression analyses were carried out with English grade equivalent, English Writing Quality, and English Spelling % Score as dependent variables in order to compare the present results with those of Cummins et al (1984). Table 5, p. 198, presents the summary regression analyses data for these variables with LOR, JAT, AOA and Age as predictors. Predictors were entered in a fixed order for all analyses. The analyses show that students' grade equivalent in English is strongly predicted by both length of residence in Canada and variables representing students' cognitive/academic development (Japanese T-scores, Age of Arrival, and Age). These latter variables account for an additional 20% of the variance in English grade equivalent after length of residence has been entered into the equation. Minimal variance is accounted for in the English writing variables, however. Although the Quality Rating and Spelling % Score do correlate significantly with both English Grade Equivalent and English T-Score (r = .29 and .14 for Quality Rating and r = .32 and .34 for Spelling % respectively), they appear to be considerably less affected by either amount of exposure to English (as represented by LOR) or students' prior cognitive/academic development in Japanese (as represented by JAT, AOA and Age) than is the case for the English reading standardized test scores.

To what extent is English writing performance related to Japanese writing performance? Expressed differently, to what extent is there interdependence of writing skills across languages as different in their writing conventions as English and Japanese? This question was addressed by regressing English Grade Equivalent, English Writing Quality and English Writing Spelling against Length of Residence (LOR), Japanese T-Score, Age of Arrival (AOA), Age, and 27 Japanese writing variables derived from the letter measure (see Table 6, p. 198). The first three variables were entered in a fixed order followed by stepwise selection of the Japanese writing variables to the specified tolerance levels (partial r significant at $p < .05$). The extent to which Japanese writing variables enter the equation after the Japanese T-Score, AOA and Age, can be interpreted as specific writing-related interdependence that is independent of general cognitive/academic skills. On the other hand, partial correlations between the dependent and independent variables controlling for LOR (i.e. after LOR has been entered into the equation) indicate the relationship across languages that is inclusive of general cognitive/academic skills.

The N is considerably reduced for these analyses in comparison to those reported previously due to missing data for some of the Japanese variables.

For English Grade Equivalent (EGE), LOR and the cognitive/academic block of variables (JAT, AOA, Age) account for 68% of the variance (adjusted R square). Two additional writing variables enter the equation - Number of errors in Hiragana (JC2) and Number of Chinese characters (JC3). Both of these variables have significant beta weights in the final regression analysis. In addition to these two variables, Organization of Ideas (JC9) has a partial correlation of .46 ($p < .0001$) with EGE. Five additional Japanese writing variables have significant partial correlations of less than .40 with EGE, suggesting a relatively strong cross-lingual relationship of general academic skills.

Length of residence accounts for 12% of the variance in English Writing Quality while the cognitive/academic block adds an additional 13%, primarily due to the strong effect of AOA. The only additional variable to enter the equation is Overuse of "I" (JC25). This variable did not quite attain significance in the partial correlations but several other Japanese writing variables did relate significantly ($p < .05$) with English Writing Quality. Specifically, these were Number of Letters Used (JC1) (partial r = .29), Number of Chinese Characters (JC3) (partial r = .29), Errors in Chinese Characters (JC4) (partial r = .29) and Level of Syntactic Structure Sophistication (JC7) (partial r = .26).

For English Spelling, AOA is again the major variable in the 23% explained variance accounted for by LOR and the cognitive/academic block. Japanese Spelling (Katakana) (JC5) adds 13% to the explained variance and Appropriateness of Topic (JC14) adds a further 7% to bring the total explained variance to 30 %. In addition to these two variables Number of Chinese Characters (JC3) (partial r = .33), Appropriateness of Expressions (JC12) (partial r = .32), Use of Letter Writing Formulae

(JC13) (partial $r = .31$), and Appropriateness of Topic (JC14) (partial $r = .31$) correlated at significance levels of $p < .01$ with English Writing Spelling. Number of Letters Used (JC1) and Number of Errors Hiragana (JC2) had partial correlations of $p < .05$).

These data suggest that in addition to cross-lingual writing relationships that are mediated by general cognitive/academic skills, there are some writing-specific cross-lingual relationships that are largely inde[illegible]dent of general cognitive/academic skills. For example, the relationship between Japanese Katakana Spelling (JC5) and English Spelling is virtually unaffected by the entry of JAT, AOA and Age into the regression equation, falling only from a partial r of .44 to .40.

4. CONCLUSION

The results of the different analyses provide a consistent picture in relation to the questions posed at the beginning of the chapter. First, although the sample as a whole performs close to the mean in Japanese reading skills, there is a clear negative relationship between length of time in Canada and students' Japanese reading proficiency. The negative effect of length of residence on Japanese writing, however, appears minimal. Age of arrival in Canada appears to be a more potent force in predicting maintenance of Japanese writing skills than length of residence. Similarly for Japanese reading, the older students are when they come to Canada, the better prospects they have for strong continued development of Japanese reading skills. This effect is not entirely due to the fact that students who arrive at older ages tend to have spent less time away from Japan, since the partial correlation between Age of Arrival and Japanese T-Score remains significant even when Length of Residence is controlled.

It appears that students require about four years' length of residence, on the average, to attain grade norms in English reading skills. There appears to be some tendency for students who arrive at the age of 6-7 to make somewhat more rapid progress towards grade norms than those who arrive at older ages. A similar tendency was evident in the Cummins (1981) results.

When length of residence is controlled, a significant relationship emerges between Japanese reading skills and English reading. Students' age of arrival in Canada (AOA) is also strongly related to English reading (controlling for LOR), suggesting the influence of general cognitive maturity in mediating the cross-lingual relationship of cognitive/academic skills. General cognitive maturity, however, cannot account fully for the interdependence of reading skills across languages since significant relationships across languages were found for reading T-scores, in which the effects of age have been removed.

Writing performance was less closely related across languages than was the case for reading. This may be partly a function of the different types of measures used in each case (standardized reading tests v. non-standardized writing tasks). However, consistent significant relationships were obtained betewen Japanese writing and both English reading and writing measures. For some variables (e.g. Spelling) there was strong evidence of a specific cross-lingual relationship that was not mediated by more general cognitive/academic proficiencies.

In general, the data are consistent with previous studies in supporting the interdependence of cognitive/academic skills across languages. They also suggest that at least four years is required for students from highly educated backgrounds to attain grade norms on English academic tasks and that continued development of L1 academic skills to a high level (i.e. that of students in the home country) is a formidable task for students who arrive in the host country at an early age (particularly prior to formal schooling) but is considerably less problematic for students who arrive after several years of schooling in their home country.

References

Cummins, J. Age on arrival and immigrant second language learning in Canada: A reassessment. **Applied Linguistics**, 1981, 2, 132-149.

Cummins, J. **Bilingualism and Special Education: Issues in Assessment and Pedagogy.** Clevedon, England: Multilingual Matters, 1984.

Cummins, J., Swain, M., Nakajima, K., Handscombe, J., Green, D. and Tran, C. Linguistic interdependence among Japanese and Vietnamese immigrant students. In C. Rivera (ed.) **Communicative Competence Approaches to Language Proficiency Assessment: Research and Application.** Clevedon, England: Multilingual Matters, 1984.

Cummins, J. and Swain, M. **Bilingualism in Education: Aspects of Theory, Research and Practice.** London: Longman, 1986.

Harley, B. **Age in Second Language Acquisition.** Clevedon, England: Multilingual Matters, 1986.

Table 1

Means and Standard Deviations for English Reading and Writing Variables, Japanese Reading, Age of Arrival, and Length of Residence N = 226

	Mean	Standard Deviation
English Reading T-Score	41.4	16.8
English Reading Grade Equivalent	4.4	3.1
English Reading Percentile	34.9	30.4
Japanese Reading T-Score	48.4	10.4
English Writing (V1) #words in T-units	109.3	167.0
English Writing (V2) #words in mazes	35.6	185.0
English Writing (V3) #T-units	42.8	171.4
English Writing (V4) #mazes	31.1	173.4
English Writing (V5) #spelling errors	34.2	172.9
English Writing (V6) #other errors	34.6	172.8
English Writing (V7) quality rating	4.0	1.5
English Writing (V8) organization	4.0	1.4
English Writing (V9) style	4.1	1.5
English Writing (V10) interest	2.6	1.4
English Writing (V11) spelling errors %	91.7	17.7
English Writing (V12) other errors %	91.6	17.5
English Writing (V13) residual %	83.3	34.2
Age of Arrival (months)	78.4	42.8
Length of Residence (months)	45.8	39.5

Table 2

Correlations between Age of Arrival, Japanese T-Score, and English Reading and Writing Variables

	Zero Order Partials			Partials Controlling for LOR	
	AOA	JAT	LOR	AOA	JAT
E T-Score	-48*	-06	53*	-08	19*
E Grade Equivalent	-27*	-03	60*	51*	28*
# Words in T-Units	05	-02	-04	03	-04
# Words in Mazes	12	02	-17*	-05	-05
# T-Units	09	-00	-16*	-08	-07
# Mazes	09	00	-17*	-09	-08
# Spelling Errors	09	-00	-17*	-10	-08
# Other Errors	10	00	-17*	-09	-08
Quality Rating	-06	-10	21*	22*	01
Organization	-07	-07	21*	21*	01
Style	-03	-05	15	18*	-01
Interest	-03	-05	05	02	-04
Spelling Error % Score	03	04	18*	22*	12
Other Error % Score	-14	-01	25*	13	10
Residual % Score	-09	02	22*	18*	11
Age of Arrival		41*	-84*		14
Japanese T-Score	41*		-40*	14	

* $p < .01$

Table 3

Cross-tabulation of English T-Scores by Age of Arrival and Length of Residence

Age of Arrival	Length of Residence			Row Total
	0-35 months	36-48 months	49 + months	
0 - 5 years	51.50^{X} 2^{Y} 9.19^{Z}	39.86 7 6.36	52.16 68 8.53	51.03 77 9.01
6 - 7 years	40.85 26 15.13	48.80 15 6.16	53.80 10 9.25	45.73 51 12.99
8 - 9 years	31.44 50 17.35	42.69 13 15.69	53.57 7 10.56	35.74 70 17.92
10 - 11 years	27.33 30 19.70	38.67 3 10.79	.00 0 .00	28.36 33 19.23
12 + years	28.71 24 18.67	.00 0 .00	.00 0 .00	28.71 24 18.67
Column Total	32.17 132 18.67	44.26 38 11.06	52.47 85 8.70	40.74 255 17.33

Number of missing observations = 18

X = mean
Y = no. of students
Z = standard deviation

Table 4

Cross-tabulation of English Grade Equivalent Score by Age of Arrival and Length of Residence

Age of Arrival	Length of Residence			Row Total
	0-35 months	36-48 months	49 + months	
0 - 5 years	2.050[X] 2[Y] .636[Z]	2.129 7 .783	5.775 68 3.152	5.347 77 3.197
6 - 7 years	2.592 26 1.487	4.493 15 1.809	7.820 10 2.487	4.176 51 2.669
8 - 9 years	2.422 49 1.817	5.715 13 2.662	8.929 7 2.546	3.703 69 2.986
10 - 11 years	2.737 30 2.620	5.700 3 3.081	.000 0 .000	3.006 33 2.750
12 + years	3.938 24 3.052	.000 0 .000	.000 0 .000	3.938 24 3.052
Column Total	2.800 131 2.267	4.571 38 2.417	6.275 85 3.182	4.228 254 3.056

Number of missing observations = 19

X = mean
Y = no. of students
Z = standard deviation

Table 5

Regression Analyses with English Reading Grade Equivalent and English Writing Quality and Spelling as Dependent Variables (N = 226)[1]

	E Grade Equivalent		E Writing Quality		E Writing Spelling	
	Adj R sq	Beta	Adj R sq	Beta	Adj R sq	Beta
LOR	35	1.43*	04	.51	03	.39
JAT	40	.18*	04	-.06	04	.09
AOA	55	.90*	08	.39	07	.22
AGE	55	-.11	08	.01	07	.09

[1] The adjusted R squared represents the cumulative percentage of variance explained by the independent variables. The Beta weights and significance levels represent the Beta weights for the final regression equation after all independent variables have been entered.

*p < .01

Table 6

Regression Analyses with English Grade Equivalent, English Writing Quality, English Writing Spelling as Dependent Variables and LOR, JAT, AOA, Age, and Japanese Writ ing Variables as Independent Variables (N = 70)[1]

	E Grade Equivalent		E Writing Quality		E Writing Spelling	
	Adj R sq	Beta	Adj R sq	Beta	Adj R sq	Beta
LOR	45	1.18*	12	1.02*	00	.62*
JAT	47	.11	11	.03	00	-07
AOA	68	.58*	25	.74*	10	56*
AGE	(68)		(25)		(10)	
#Errors Hiragana (JC2)	70	-.18*				
#Chinese Chars (JCE3)	72	.17*				
Overuse of "I" (JC25)			29	.22*		
Spelling Katakana (JC5)					23	34*
Appropriateness Topic (JC14)					30	29*

[1] The adjusted R squared represents the cumulative percentage of variance explained by the independent variables. The Beta weights and significance levels represent the Beta weights for the final regression equation after all independent variables have been entered.

*p < .05

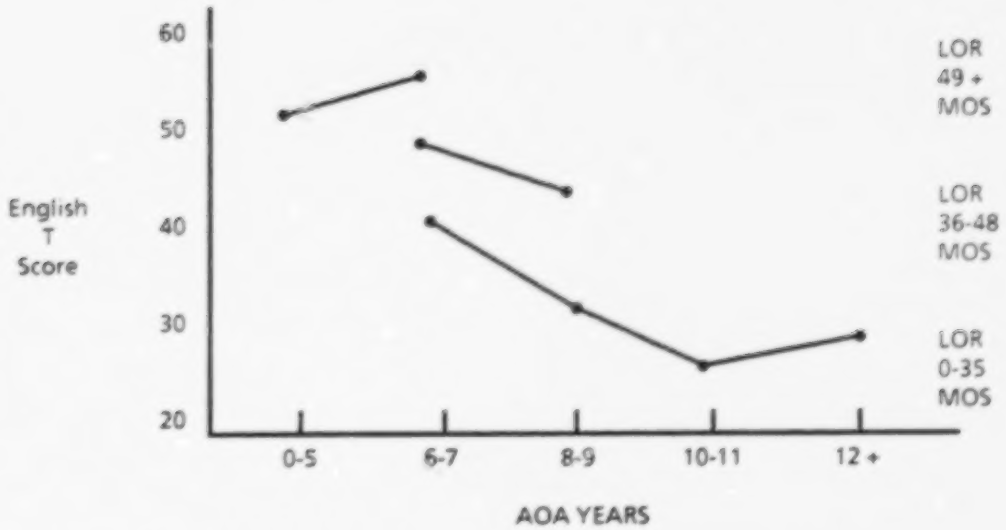


Figure 1: The Relationship between English T-Score and Length of Residence (LOR) and Age of Arrival (AOA)

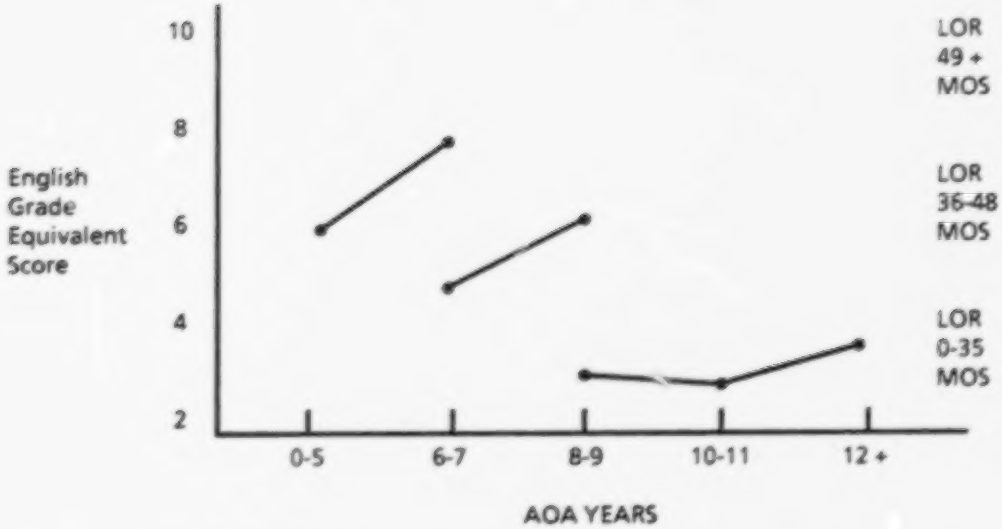
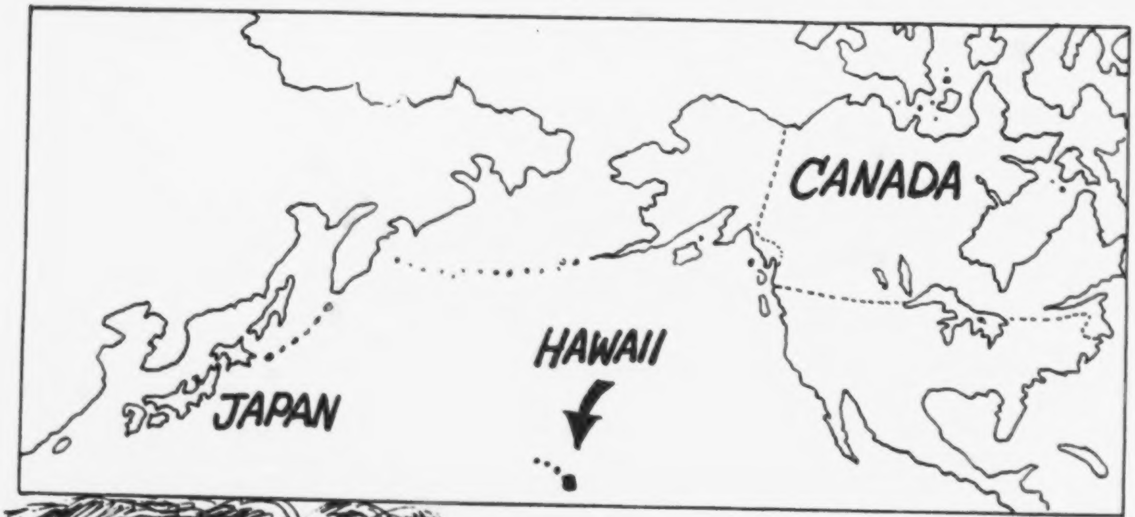


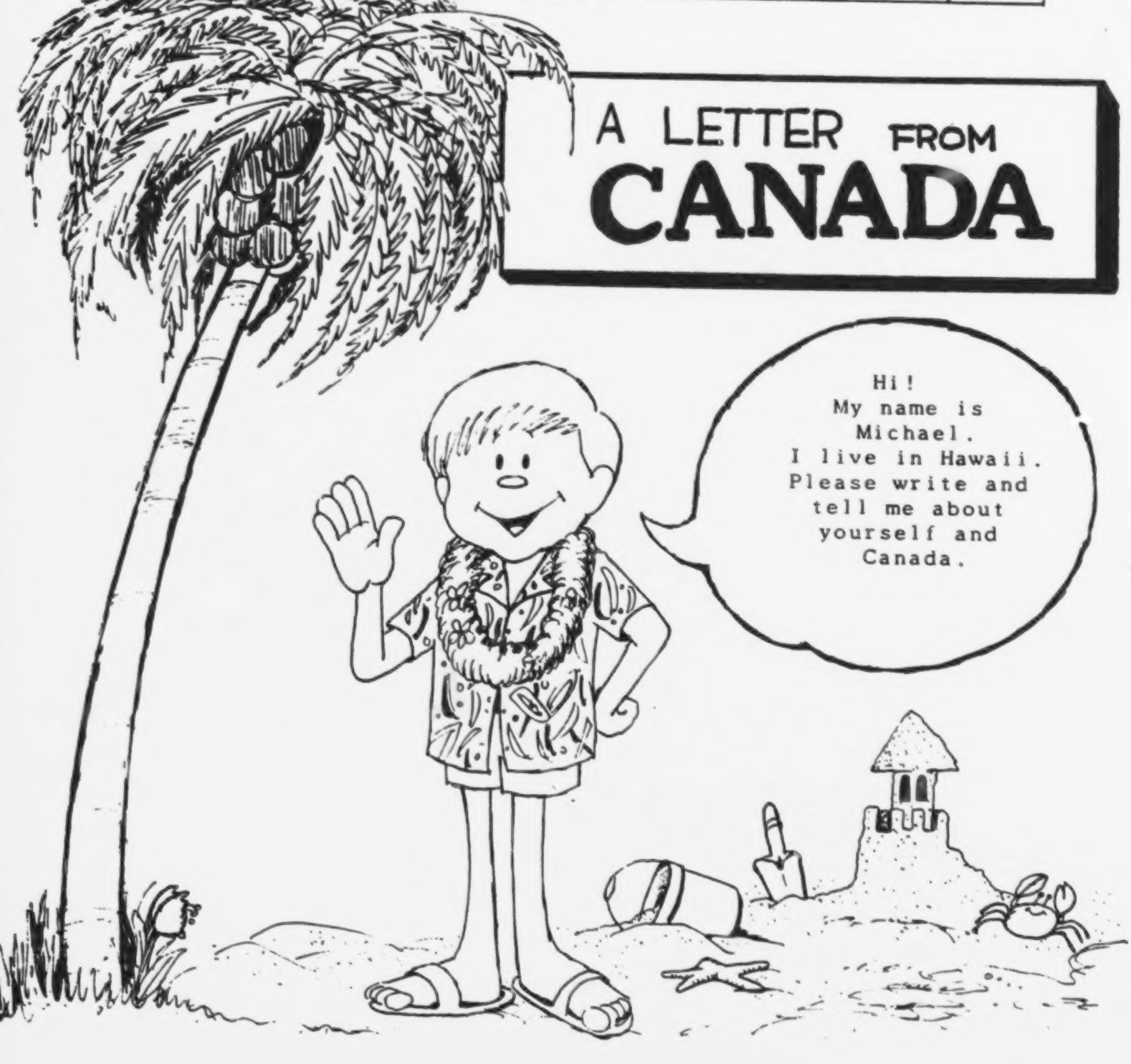
Figure 2: The Relationship between English Grade Equivalent Score and Length of Residence (LOR) and Age of Arrival (AOA)

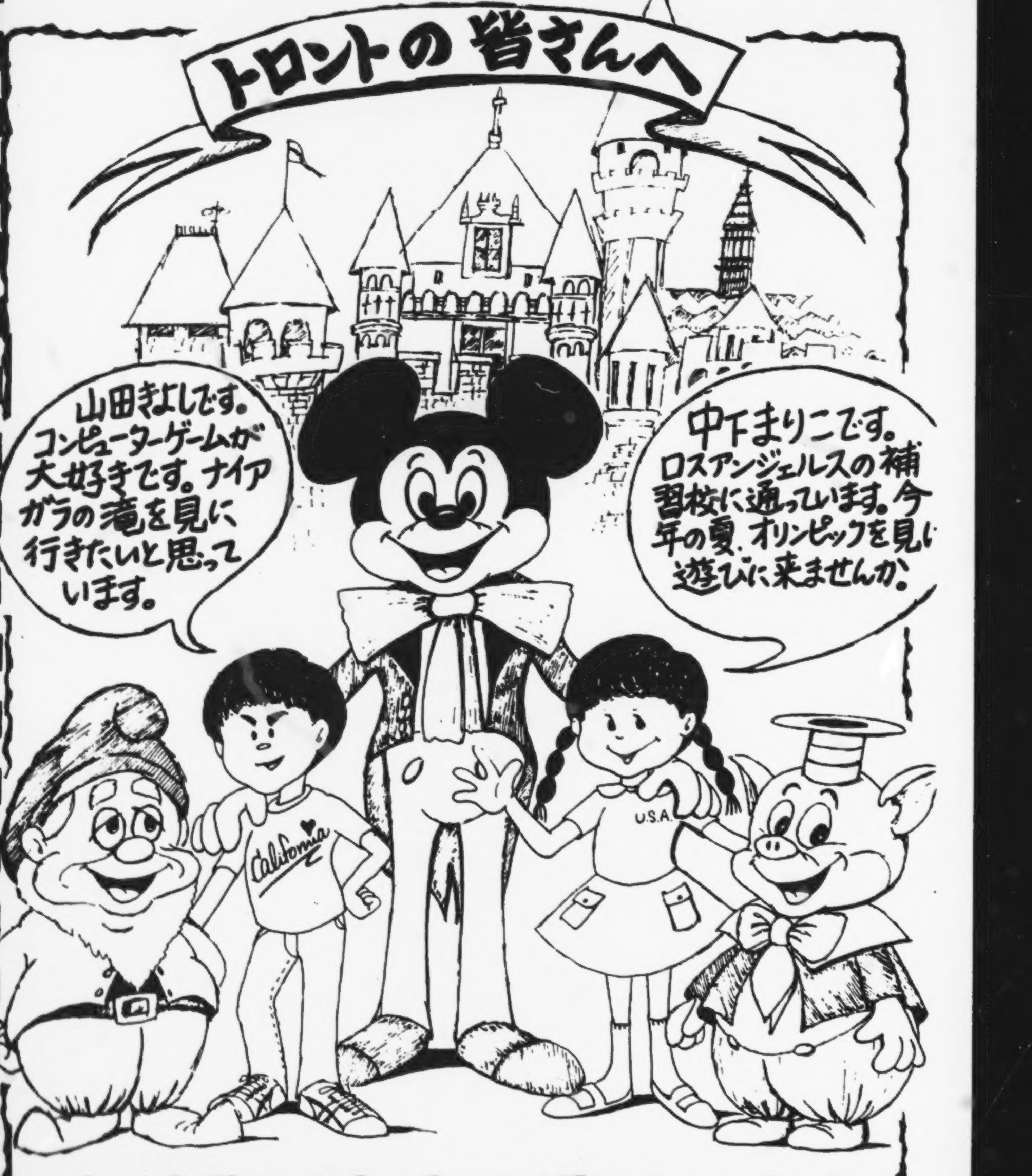
Appendix A

English and Japanese Letter Writing Tasks


CANADA
HAWAII
JAPAN

A LETTER FROM CANADA


Hi! My name is Michael. I live in Hawaii. Please write and tell me about yourself and Canada.


トロントの皆さんへ
山田きよしです。コンピューターゲームが大好きです。ナイアガラの滝を見に、行きたいと思っています。
中下まりこです。ロスアンジェルスの補習校に通っています。今年の夏、オリンピックを見に遊びに来ませんか。
California
U.S.A.

Chapter 10

THE RELATIONSHIP BETWEEN STARTING AGE AND ORAL SECOND LANGUAGE PROFICIENCY IN THREE GROUPS OF CLASSROOM LEARNERS

Birgit Harley

1. INTRODUCTION

In this study, the oral second language (L2) proficiency in French of three groups of English-speaking classroom learners is compared. One group had begun intensive exposure to French at an early age, in kindergarten, while the other two groups had started intensive exposure several years later, in grade 7. The comparison, involving grammatical, discourse, and sociolinguistic aspects of the students' L2 communicative competence, takes place at the grade 10 level, when the students are close to 16 years old.

The study builds on previous research conducted in the context of several different studies: (a) by Harley (1986), who compared the oral control of the French verb system among early partial and late immersion students in grades 9 and 10; and (b) by Lapkin and Swain (Lapkin and Swain 1984a,b; Appendix C to Chapter 5), whose investigation of the French proficiency of early total, late extended, and late immersion students in grade 10 has included analyses of their oral L2 fluency and sociolinguistic skills.

The purpose of the present study is to determine whether there are specific long-term advantages in oral L2 proficiency that can be associated with intensive L2 exposure at an early age in a total French immersion classroom setting. Studies comparing early total and late immersion students at the secondary level carried out in various Canadian centres (e.g. Genesee 1983, Morrison and Pawley 1986, Swain and Lapkin 1986, Chapter 5 Appendix C), have produced mixed findings. In Ontario, early total immersion students have maintained advantages in speaking skills over late immersion and extended French students at the secondary level (Morrison and Pawley 1986, Swain and Lapkin 1986). In Quebec, however, where a less substantial follow-up program to early immersion has been provided, early immersion students have not consistently outperformed late immersion students in oral skills at the secondary school level (Genesee 1983). Similarly, the early partial immersion students studied by Harley (1986) were ahead of late immersion students on some, but not all, features of the French verb system.

In line with other studies comparing different starting ages for school-based L2 programs, these findings indicate that time, or accumulated hours of classroom L2 exposure, is no firm guarantee of greater long-term success for those with an earlier start. An analysis of theoretical perspectives and empirical findings on the age issue in L2 acquisition (Harley 1986) suggests that maturational and environmental variables interact in complex ways to determine proficiency outcomes. While older learners may have cognitive advantages over younger learners which enable them to make faster initial progress in sorting out some of the complexities of a new language system, an early start offers the potential over time for a much greater quantity of appropriate L2 input and opportunities for meaningful, sustained L2 production incorporating a wide range of L2 features. The extent to which that potential is realized in different classroom contexts may be a key factor influencing the comparative results that are found.

In the present study, it may be hypothesized that any differences found between early total immersion and late-entry students (in immersion or extended French programs) at the secondary level will generally favour the early immersion students, on the assumption that they will have had considerably more relevant L2 input and experience in using French for a wide range of oral communicative purposes over the course of their program. One exception to this expected pattern of findings may be in the sociolinguistic domain, where early immersion students at the secondary level have previously been found lagging in the appropriate use of formal second person forms to express politeness (Appendix C to Chapter 5). As is clear from the classroom observations reported in Chapter 5, early immersion students, at least in elementary school, appear to have little exposure to such formal patterns of use and are not required to produce them.

2. THE SAMPLE

The sample for the study includes three groups of learners and a group of native French speakers, for a total of 47 students. The learners were selected from the files of the "Second Language Maintenance at the Secondary School Level" Project, a research project running concurrently with the Development of Bilingual Proficiency Project at the Ontario Institute for Studies in Education (Lapkin and Swain 1984a, b). Each learner group has had a different French program background: early total immersion, late immersion, or late extended French. The first group, consisting of twelve early immersion students from programs in the Ottawa and Toronto metropolitan areas, received a half-day kindergarten in French, followed by one or two years of instruction entirely in French. Thereafter, English was gradually introduced into their program, and from grade 6 on, the students have been taking about half their subjects in French and the remainder in English. The second group of twelve late immersion students from the Ottawa area participated in a regular core French program until the end of grade 6, involving 20 minutes of French instruction per day from kindergarten to grade 5 and 40 minutes per day in grade 6. In grades 7 and 8, the students were immersed in French for 80% of their schooling, followed in grades 9 and 10 by a reduction of time devoted to French to around 40%. The third group of eleven[1] late extended French students from the Toronto area generally had 40 minutes of core French per day from grade 4 (or earlier) to grade 6, followed by 25% of the day in French in grade 7 and about 40% in later grades. It should be noted that participation in each of the three programs — early immersion, late immersion and extended French — is optional. This implies high motivation to learn French by all the learners involved in this study.

The following criteria were used in the selection of the learners: (a) they had to be chosen from among those students who had not been orally tested in the context of the Second Language Maintenance Project[2]; and (b) the three groups were to be matched as closely as possible on variables such as sex, age, IQ, and home language background. In addition to the three groups of second language learners, a norm group of twelve native French speakers attending grade 10 in a French-speaking school in a suburb of Quebec City was included in the study. This group of students had previously been used as a norm group in the study by Harley (1986).

Table 1, p. 223, displays background information about each of the four groups of students. There were fewer boys than girls in each of the learner groups (five boys in the early immersion group, and four in each of the late immersion and extended French groups), with six boys and six girls in the native speaker group. The mean age of the three learner groups was approximately 15 years, 10 months at the time of oral testing; the native speaker group was on average one month younger. Most of the early immersion students were attending school in the Ottawa area, but there were also three students from Toronto area schools. The late immersion students were all from the Ottawa area, and the extended French students from Metropolitan Toronto. The home language used most frequently by almost all the L2 learners was English, with one early immersion student indicating that he spoke English and German at home. Three early immersion students and one extended French student indicated that they also sometimes spoke French at home. There were no significant differences in IQ scores among the learner groups, who had received the Otis Lennon Mental Ability Test, Intermediate Level, in grade 9. The native speaker group had received a French version of the Advanced Level of the same test in grade 10. Scores on this test are not directly comparable to those of the three learner groups.

3. PROCEDURES

All the students took part in an oral interview. The 35 learners also received an oral sociolinguistic test.

3:1 The Interview

Each student was individually interviewed by a native French speaker from Quebec, who was one of two young women unknown to the students. For the learner groups, the interviewer was a member of the staff of the Development of Bilingual Proficiency Project, while in the case of the native speakers, who had been interviewed two years earlier, the interviewer was a former French immersion teacher. The interviews lasted up to half an hour and took place in spare classrooms or offices at school. They were tape-recorded on cassette, with a tie-clip microphone attached to the student to enhance sound quality.

Verb use. In each interview, the student was asked the same set of questions (see Appendix A, pp. 232-234), designed to provide contexts for the use of a variety of verb forms and functions in French while at the same time introducing topics of general and personal interest to the students. Subjects were asked, for example, to provide a recipe, to describe some humorous cartoons, to relate exciting episodes in their lives, and to say what they would do with a large sum of money. The questions were structured so as to provide contexts for the use of different verb tenses (imparfait, passé composé, future, conditional), verb agreement in number and person (e.g., second and third person plural forms), constructions involving the use of clitic object pronouns, and a variety of lexical verbs. The same 'guided' interview had previously been used in the study by Harley (1986). The native speaker group served to provide the necessary evidence that the questions posed by the interviewer did in fact elicit the anticipated verb forms in a natural way from native speakers of French. While each interviewer was instructed to ask the questions exactly as written on a set of index cards, she was free to react to students' responses in a natural conversational manner.

Transcription of the learners' interviews in ordinary French orthography was later carried out by two bilingual members of the project staff, one a native speaker of English and the other a native speaker of French. All transcripts were checked by a second researcher. Scoring of the learners' interviews with respect to the verb system was done from the transcripts by a near-native speaker of French, based on scoring procedures described in detail in Appendix B, pp. 235-242, and previously used by Harley (1986). In essence, the scoring consisted of assessing the use of target verb forms in the context of selected questions, to provide individual scores for specific verb variables. The way in which scores were calculated in each case is summarized in relation to the individual variables in the Analysis and Results section (4:1) below. The scores for the verb variables were checked by a second near-native French speaker for inter-rater reliability purposes. The Quebec students' interviews had already been transcribed and scored for the use of verbs in the earlier study by Harley (1986). The main scorer of the learners' interviews also carried out an inter-rater reliability check on these native-speaker scores.

Fluency. In addition to being scored for verb use, students' responses to three interview questions (Questions 9, 18, and 22) were scored for oral fluency. These questions (see Appendix A, p. 232) generally elicited sustained discourse by way of a response. The assessment of fluency focussed on the extent to which the students were able to maintain a smooth, native-like flow to their discourse, without undue recourse to hesitations or other markers of disfluency. Scoring, based mainly on Olynyk et al. (1983), pertained to:

(a) the frequency with which a student manifested disfluency markers (calculated as number of disfluencies per 100 words);

(b) the global classification of the disfluencies as either 'progressive' or 'regressive' conversational repairs; and

(c) the linguistic contexts in which disfluencies occurred.

Following Olynyk et al. (see also Lapkin and Swain 1984a, b), several types of disfluency were identified (for details, see Appendix C, pp. 243-247). Repeats of lexical items or structures and the use of hesitation phenomena such as 'uh', 'um' transitions between words were classified as progressive types of self-initiated conversational repair, which do not demand of listeners that they reorganize their expectations of what is to follow. Reliance on progressive repairs of this kind is believed to be characteristic of relatively fluent speakers of the target language (Olynyk et al. 1983). Frequent use of regressive repairs, on the other hand, which demand a revision of expectations on the part of the listener, is thought to be characteristic of less fluent speakers. Regressive repairs were defined by Olynyk et al. to consist of repair conversions or cut-offs, involving a change or break in an expression or word. In addition to these regressive types of repair, the frequency of unfilled pauses of two or more seconds in length was assessed in the present study. Such pauses were also deemed to be regressive in nature and characteristic of less fluent speech (Hulstijn 1983).

A further analysis was carried out to determine the linguistic contexts in which 'uh' and 'um' etc. transitions were used. This analysis was based on the hypothesis that fluent speakers would generally restrict such hesitation phenomena to certain contexts, such as major constituent boundaries, whereas less fluent speakers may insert them in a wider range of linguistic contexts.

Full details of the scoring criteria used, with examples of each type of disfluency and the linguistic contexts in which hesitation phenomena were tallied is provided in Appendix C, pp. 243-247. Also included in the appendix are sample scoring sheets used in tallying type and location of disfluencies (see pp. 248-250).

Scoring of fluency was done from the cassette tapes, with transcripts as additional support, by two near-native speakers of French, each of whom scored all the tapes. An initial inter-rater reliability check revealed some discrepancies in scores. Following a clarification of the scoring criteria, the tapes were rescored and each scorer rechecked the other's scores.

3:2 Sociolinguistic Oral Test

Following their individual interviews, each of the learners took a sociolinguistic oral test in French that was based on the one used with grade 6 immersion students in an earlier Development of Bilingual Proficiency Project study (see Year 2 Report, Swain 1985). For this individual test, students were shown a series of slides illustrating a variety of social situations that called upon them to make requests, offers and complaints. Each slide was accompanied by a taped explanation of the situation in French, and the student's task was to respond to the situation in a way that was sociolinguistically appropriate in French. The social status of the addressee shown on each slide was systematically varied, so that for each category of situation there was a high status variant and a low status variant. The students' responses to the situations were recorded for later scoring.

For scoring purposes, points were allotted for a variety of linguistic markers of formality used by the student in responding to each slide (see Appendix D, pp. 251-252 for details). In the high status variants of situations, use of a greater number of formal markers would be expected among native speakers of French than in the low status variants (see e.g. Year 2 Report). Final test scores were therefore based on the students' ability to adjust register appropriately between the high and the low status situational variants. For each student, this entailed subtracting the number of formal markers of each kind used in responding to the low status variants from the number of such markers used in responding to the high status variants, to produce a set of sociolinguistic 'difference' scores. The allocation of points for formal markers was done by two scorers from the tapes, one a near-native speaker and the other a native speaker of French. To determine inter-rater reliability, they both scored 20 of the 35 tapes independently (see Appendix D, Table D1, p. 253). Some disagreements on three markers: 'use of person', 'lexicon' and 'attenuation' were noted. Discrepancies were found due to misunderstanding of the criteria on the part of one scorer. This scorer's ratings on these markers were therefore revised by a third independent scorer.

4. VERB USE: ANALYSIS AND RESULTS

Two kinds of scores were calculated with respect to the use of verbs in the interview. Total scores represented a simple count of target verb items produced in the context of specific questions. Percentage scores represented total scores divided by the sum of items produced in the given contexts, as defined briefly below and explained in more detail in Appendix B, pp. 235-242. In the case of two variables, clitic pronoun complements and verb lexis, total scores only were calculated, consisting respectively of the number of non-subject clitic pronouns and the number of different verb types produced in the context of a set of picture descriptions. Inter-rater reliabilities were based, except for clitic pronouns and lexis, on the percentage scores. As Table 2, p. 224, indicates, the reliabilities were generally high, with r ranging from .73 to .99 on specific variables.

The prediction was that the early immersion students would generally be more native-like in their use of French verbs than the other learner groups, except in the case of the sociolinguistically relevant use of polite second person (plural) verb forms, where the opposite prediction was made. These predictions were tested on the interview data by comparing the groups on the non-parametric Mann-Whitney U Test. This test was used in preference to t-tests since the distribution of scores on some variables was skewed owing for example to several zero or 100% scores on particular variables. Scores are reported on Tables 3 and 4 (pp. 225 and 226) as both means and medians, the latter calculated by means of SPSS-X.[3] On Tables 5 and 6 (pp. 227 and 228) the range of scores in each of the groups is displayed. The results for each variable are presented below.

4:1 Deictic Time Distinctions

Several questions in the interviews were designed to tap the learners' ability to make deictic time distinctions in the verb: that is, to situate the time of an event or process relative to the moment of speaking. In the present study, the analysis focussed on the students' use of future tenses in responding to questions 10 and 19, which involved the prediction of events that had not yet taken place, and on their use of past tenses in responding to questions 23, 34, 34a, 34b and 35, which called for the narration of events that had occurred in the past (for the content of the questions concerned, see Appendix A, pp. 232-234).

The future. Two future variables were analysed in order to determine the extent to which the learners were (a) expressing future time in the verb, and (b) using the periphrastic future with comparable frequency to the native speakers. The total score for the first future variable consisted of all instances of future tenses, as well as conditional and elliptical infinitive forms with future time reference. The number of contexts for future time reference provided the percentage base. As Table 3, p. 225, shows, both the early and the late immersion students produced on average similar numbers of non-present tense verb forms with future reference, and did not differ significantly from the native speakers. The extended French students, on the other hand, who had started their intensive program in grade 7 but had devoted less total time to French than the late immersion students, produced significantly fewer such verb forms than the native speakers ($p < .01$). In their percentage scores on this first future variable, it can be seen from Table 4, p. 226, that only the early immersion students, with a median score of 91.7%, were not significantly different from the native speakers with a median of 100%. At the same time the early immersion students' percentage scores were significantly higher than those of the extended French students with a median of 50%.

Periphrastic future. This initial positive finding for early immersion students has to be tempered, however, in light of the findings on the second future variable, the periphrastic future. The total score for this second variable consisted exclusively of periphrastic future forms (ALLER + infinitive), while the percentage base remained the same as for the first future variable. Tables 3 and 4 show that, compared with the native speakers, all three learner groups were using similarly few periphrastic future forms, and that all three differed significantly from the native speakers in total and percentage scores. A comparison of scores on the two future variables indicates that whereas, for the native speakers, the great majority of verbs with future reference occurred in the periphrastic future, for the various learner groups this was apparently not the case. Inspection of the scoring sheets reveals that the learner groups all tended to rely more heavily on the use of elliptical infinitive forms. (Examples of such forms are provided in Appendix B, p. 238). In comparing the way in which the learners and the native speakers expressed future time, therefore, it appears that contrary to the initial prediction, the early immersion students were not more native-like than the other learner groups.

Past time. The total score for this variable consisted of the number of accurately used past tense forms (passé composé, imparfait, etc.) in referring to past time, not including use of the historic present, or auxiliary (AVOIR instead of ETRE) and past participle errors (see Appendix B, p. 237). Errors in number and gender agreement were disregarded. The percentage base consisted of all verb forms with past time reference, including present tense forms and auxiliary and past participle errors. As can be seen in Tables 3 and 4 (pp. 225 and 226), all three learner groups made significantly less use of accurate past tense forms than the native speakers in referring to past time. There were, however, no significant differences between the early immersion students and the other learner groups, once again disconfirming the prediction made, although a tendency may be noted for the percentage scores of the late extended French students to be lower than those of both immersion groups.

4:2 Aspectual Distinctions

As well as providing contexts for past narratives, the interview was designed to elicit aspectual distinctions in the past. Questions 13 and 39, for example, provided contexts for the expression of past actions in progress, while question 28 was designed to elicit reference to habitual or repeated past actions (see Appendix A, pp. 232-234). In French, such aspectual distinctions are generally expressed in the imparfait.

The imparfait. Total scores for the use of the imparfait in the context of questions 13, 28 and 39 consisted, for each student, of a simple count of uses of this verb form. As Table 3 shows, all three learner groups differed significantly in total scores from the native speakers, who made on average more than three times as much use of the imparfait as any of the learner groups. At the same time, the early immersion students produced on average significantly more imparfait forms in the given question contexts than did the late immersion and extended French groups. The late immersion students, in turn, showed a marginally significant tendency to produce more imparfait forms than the extended French students who had had less exposure to French in their program (1-tailed $p < .05$). The percentage scores for the imparfait, displayed on Table 4, confirm that, of the learner groups, the early immersion students were on average most like the native speaker group, from whom they did not differ significantly. These percentage scores were calculated by dividing total scores by all verb forms produced in answering the given questions, except when set expressions were used with obvious non-past reference (see Appendix B, p. 238). This method of calculation was used in preference to using 'obligatory' contexts as a percentage base owing to the difficulty of establishing the intended meaning of the students in all instances and thus determining the obligatory contexts for use of the imparfait. As can be seen from the percentage scores on Table 4, not even the native speakers (with a median score of 70.7%) used the imparfait all the time in responding to the relevant questions. With their median percentage score of 63.3%, the early immersion students were not only close to the native speakers but highly significantly superior to the late immersion students with a median score of 13.4%. The late immersion students, again, were significantly ahead of the extended French students with a median score of 0%.

A closer analysis of the use of the imparfait in the progressive (questions 13 and 39) and habitual past (question 28) contexts shows that whereas the native speakers made on average most use of the imparfait in the habitual context of question 28 (on average over 4 instances per student compared with less than 3 instances per student in the combined contexts of questions 13 and 39), the early immersion students made comparatively much less frequent use of the imparfait in the habitual past context (on average less than 0.4 instances per student in the context of question 28 versus more than 2 instances per student in the combined progressive contexts of questions 13 and 39). These findings suggest that the early immersion students, despite the similarity of their overall scores on the imparfait to those of the native speakers, were not yet regularly using the imparfait in the context of habitual past actions.

4:3 Hypothetical Events

Two questions in the interview were designed to elicit use of conditional verb forms in the expression of remotely possible hypothetical events (questions 36 and 37). A total score for each student consisted of all accurately realized conditional forms used in hypothetical contexts. The percentage base consisted of obligatory contexts for the use of the conditional. All three learner groups differed highly significantly from the native speakers on both total and percentage scores (see Tables 3 and 4). In keeping with the initial prediction, however, the early immersion students in turn produced significantly more conditional forms than the other learner groups (Table 3), and in significantly greater proportion to the obligatory contexts supplied (Table 4). In fact, whereas the early immersion students had a median percentage score of 55%, the late immersion and extended French students' median percentage scores were both at 0%, indicating a virtually complete lack of productive control of the conditional verb form in a hypothetical context.

4:4 Number Distinctions

Several questions in the interview were designed to elicit the use of plural number agreement in verb forms. Responses to these questions are considered below in relation to first, second, and third person contexts.

First person plural. Questions 17, 17a, 23, 23a and 26 were designed to elicit talk about the actions of the speaker (the interviewee) together with others (i.e. the first person plural).. A total score for this first person plural variable was arrived at by summing instances of verb forms agreeing accurately with the subject pronouns nous and on. Verb errors in, for example, tense, auxiliary and past participle were ignored. Percentage scores were calculated by dividing the total scores by the number of contexts produced for referring to the actions or states of the speaker plus others. The percentage base thus included elliptical infinitives and errors in number agreement. A comparison of the total and percentage scores of the various groups reveals that, contrary to the prediction made, the early immersion students were, of the three learner groups, least like the native speakers on this variable. In total scores, all three learner groups were substantially and significantly lower than the native speakers (see Table 3), but the early immersion students also showed a near-significant tendency to use fewer instances of such first person forms than the late immersion and extended French groups (1-tailed $p < .05$). In percentage scores, only the early immersion students differed significantly from the native speakers and also scored significantly lower than both the late immersion and extended French groups.

***On* as first person plural.** Since the native speakers made exclusive use of *on* + verb stem forms in the first person plural context (see identical native speaker scores on Tables 3 and 4 for the 'first person plural' and '*on* first person plural' variables), a second analysis was carried out to see how native-like the various immersion groups were in this regard. It might be expected that the early immersion students, with more exposure to informal colloquial French in the context of their program, would be more like the native speakers than the other learner groups. This did not appear to be the case, however. A substantial portion of the first person plural forms used by the early immersion students took other forms, including the formal *nous* + *-ons*, more appropriate in formal, written French than in oral conversation. Indeed, the late immersion students appeared near-significantly more likely (1-tailed $p < .05$) than the early immersion students to use *on* plus verb stem in the given contexts (see Table 4).

Second person plural. A total score for this variable consisted of all second person plural verb forms agreeing correctly with the subject pronoun produced in response to question 40 on the interview. A percentage score was calculated by dividing the total score by the number of obligatory second person (singular or plural) contexts produced. In some instances, students resorted to indirect speech rather than direct speech, thus bypassing the need to use second person forms. This was particularly the case among the native speakers (see Table 4). Where no contexts for second person forms were produced, percentage scores were considered to be missing. Table 4 shows that relative to the native speakers, two of the learner groups — early immersion and extended French —displayed a near-significant tendency towards lower percentage scores (1-tailed $p < .05$). If there had been a greater native speaker N (it was only 3), there would doubtless have been a higher level of significance. The late immersion students, on the other hand, who showed a non-significant tendency toward greater proportional use of second person plural forms than either of the other learner groups (see Table 4), did not differ significantly from the native speakers. As in the case of first person plural forms, therefore, the prediction that early immersion students would be more native-like than the late immersion and extended French students was not maintained.

Polite second person. In responding to question 11 in the interview, students had to role-play a polite request to a stranger on the street. The native-speakers in this situation almost all used the pronoun *vous* and associated verb forms. Not one of them addressed the stranger as *tu*. For this variable it had been predicted on the basis of prior findings (Swain and Lapkin in press) that the early immersion students would be *less* native-like than the other learner groups. This prediction was upheld. Only the early immersion group was significantly lower than the native speakers, both on total use of second person forms agreeing with *vous* and on percentage use of these forms in second person contexts (Tables 3 and 4). With their median percentage score of 50%, the early immersion students also showed a near-significant tendency (1-tailed $p < .05$) to less proportionate use of the appropriate verb forms in second person contexts than the other two learner groups, each of which had a median percentage score of 100% (see Table 4).

Third person plural. Several questions on the interview involved the description of picture cartoons in which there was more than one protagonist (questions 29, 29b, 30, 30b, 32 and 33), providing contexts for the use of third person plural verb forms. The total score for this variable consisted of the number of phonologically distinct third person verb forms produced in the picture descriptions, and the percentage score consisted of the proportion of such forms produced accurately in obligatory context. The results of the group comparisons on this variable failed to demonstrate the predicted advantage for the early immersion students over the other learner groups. There were no significant differences between the early immersion students and the late immersion and extended French students, while average total and percentage scores of the native speaker group were generally significantly higher than those of all the learner groups. Only the total score of the late immersion group did not differ significantly from that of the native speakers. Notwithstanding these results, there is a slight indication from the median percentage score of the early immersion students (73.3%) that they tended to be more native-like than the other learner groups with medians of 50%.

In general, however, the findings on number agreement in the verb show the early immersion students either at a disadvantage to the other learner groups or not further ahead in relation to the native speakers, thus disconfirming the prediction that early immersion students would be more native-like on most of these variables.

4:5 Clitic Pronoun Complements

The preceding verb variables concentrate on the various learner groups' use of morphological distinctions in verb inflections. The present variable, clitic pronoun complements, concentrates on the learners' use of verb phrase syntax, with a focus on the order of non-subject pronouns vis-à-vis the verb. A total score for this variable consisted of all non-reflexive, pre-verbal pronoun complements that were accurate in case, but not necessarily in gender, and that were produced in response to questions 5, 9, 10, 12, 31, 31a, 33 and 33a -- questions that tended to give rise to the expression of direct and indirect objects. There was no percentage score for this variable. As indicated in Table 3, there is some evidence to support the hypothesis that the early immersion students would be more native-like in their use of these pronouns than the late immersion and extended French students. The early immersion students produced significantly more such pronoun complements than the extended French students, and whereas both the late immersion and extended French students, with median scores of 2.0 and 1.0 respectively, had significantly lower total scores than the native speakers (with a median score of 5.0), the scores of the early immersion students (median = 4.5) differed from those of the native speakers at only a marginal level of significance (1-tailed $p < .05$).

4:6 Pronominal Verbs

Three questions in the interview, 9, 10 and 12, provided contexts for the use of pronominal verbs in the expression of actions such as getting up (se lever), dressing (s'habiller), bathing (se baigner), and going to bed (se coucher). A total score consisted of all pronominal verbs used in response to these questions that were accurate with respect to use of reflexive pronouns and, where relevant, auxiliary être. The base for the percentage score consisted of all contexts provided for the use of pronominal verbs in responding to the same three questions. All three learner groups had significantly lower total and percentage scores on this variable than did the native speakers, and there were no significant differences among the learner groups (see Tables 3 and 4). A tendency for the early immersion students to do better on pronominal verbs than the late immersion and extended French students is, however, indicated by the smaller apparent difference in total and percentage scores between the early immersion and native speaker groups than between the other learner groups and the native speakers. This tendency provides only very tentative evidence in support of the overall prediction of more native-like use of the verb system by the early immersion students than by the other learner groups.

4:7 Lexis

The final verb variable examined in the present study was lexical in nature. A total score only was calculated, based on the number of different lexical verbs (verb types) produced in the picture descriptions elicited by questions 9, 10, 29, 30, 31, 32 and 33. Grammatical errors in the verb were disregarded in arriving at a score for each student (for further details, see Appendix B, p. 241-242). Contrary to prediction, the early immersion students were not ahead of the other learner groups in the variety of verbs that were used. Instead, a near-significant tendency may be noted for the late immersion students to produce more different verbs than the early immersion students in the given picture description contexts (1-tailed $p < .05$).

4:8 Summary of Verb Findings

The analyses of verb use in the interview setting by the three learner groups and the native speakers show that, of the learners, the early immersion students are more native-like on some variables (imparfait, conditional, use of pronoun complements in clitic position), show a slight tendency to be more native-like on others (third person plural and pronominal verbs), but are less native-like than at least the late immersion students on the various first and second person variables and possibly also on lexis, and roughly comparable to the late immersion students in the area of time distinctions. As the range of scores displayed in Tables 5 and 6 (pp. 227 and 228) demonstrates, there is considerable within-group variation on each of the variables. These findings are discussed in section 7 below.

5. FLUENCY: ANALYSIS AND RESULTS

Two aspects of fluency were compared across the three groups of learners and the native speakers: (1) the number of times that different types of disfluency were manifested, and (2) the syntactic locations of the most frequent type of disfluency, i.e., 'uh', 'um' etc. transitions (henceforth referred to as 'transition markers'). It was hypothesized that the early immersion students would emerge as the most fluent (native-like) of the learner groups.

5:1 Types of Disfluency

Table 7 (p. 229) presents group means per 100 words for each type of disfluency that was scored (see Appendix C, pp. 243-245, for definitions), together with the results of t-tests comparing the different groups. It shows that progressive types of disfluency (repeats and transition markers) were the most frequent for all groups, with the transition type representing by far the most common disfluency phenomenon. All three learner groups produced significantly more repeats and transition markers than the native speakers, but there were no significant differences among the learner groups on these progressive disfluencies.

With respect to regressive types of disfluency, involving unfilled pauses, repair conversions and cut-offs, it is evident from Table 7 that the native speakers in general demonstrated far fewer instances of these disfluencies than the learner groups. The differences between the learners and the native speakers were in most cases statistically significant, particularly with respect to unfilled pauses and repair conversions. In the area of cut offs, there was some evidence that the early immersion students were more native-like than the other learner groups. Thus, while all three learner groups made significantly more cut offs within a word than did the native speakers ($\overline{X} = 0.17$), the early immersion students, with a mean of 0.41, made significantly fewer within-word cut offs than the extended French students ($\overline{X} = 1.40$, $p < .01$) and fewer at a marginal level of significance ($p < .10$) than the late immersion students too ($\overline{X} = 1.49$). Cut offs involving a sound change did not differ significantly among the learner groups (see Table 7), but the early immersion students ($\overline{X} = 0.57$) differed only at a marginal level of significance on this variable from the native speakers ($\overline{X} = 0.10$), while the scores of the late immersion ($\overline{X} = 0.89$) and extended French students ($\overline{X} = 0.87$) differed significantly from those of the native speakers at the .05 and .01 levels respectively. On the third type of cut off, involving a structural change in an expression or sentence, only the extended French students ($\overline{X} = 1.97$) differed significantly from the native speakers ($\overline{X} = 0.73$), while there was no significant difference for the early immersion students ($\overline{X} = 1.77$) and a marginally significant difference ($p < .10$) for the late immersion students ($\overline{X} = 1.32$). When all types of cut off were considered together (see Table 7), the early immersion students with a mean of 2.75 cut offs produced significantly fewer disfluencies of this type than did the extended French students ($\overline{X} = 4.24$, $p < .05$), but did not differ significantly from the late immersion group ($\overline{X} = 3.70$).

Overall, the three learner groups produced substantially more progressive and regressive disfluencies than did the native speakers (see Table 7). At the same time, the proportion of progressive disfluencies in relation to all disfluency phenomena was, as might be expected, significantly higher for the fluent native speakers ($\overline{X} = 82\%$) than for each of the learner groups, with means of 61% for the early immersion students, 67% for the late immersion students and 61% for the extended French students. There were no significant differences in the proportion of progressive disfluencies among the learner groups, although this measure might have been anticipated to favour early immersion students, who were expected to be more fluent overall than the late immersion and extended French students.

5:2 Location of Disfluencies

The syntactic positions in which the most frequent type of disfluency — 'uh', 'um', etc. transitions — occurred were analysed and compared across groups (see Table 8, p. 230). The syntactic description, with examples, of each of the locations that was isolated is provided in Appendix C, pp. 245-247 (see also Figure C2, pp.249-250).

Table 8 shows that location 1, clause-initial position, was the syntactic location where all groups, including the native speakers, were most likely to produce transition markers. The late immersion students, with the highest mean of 4.50 for the use of transition markers in this location, were the only learner group to differ significantly on location 1 from the native speakers.

Location 2, between a noun phrase and a verb phrase (if any), was also a frequent location for transition markers in all three learner groups, although not among the native speakers (see Table 8). Each of the learner groups differed significantly from the native speakers in this location, but there were no significant differences between the early immersion students and the other learner groups.

Locations 1-6 all represent positions between clauses or phrases, i.e. major syntactic constituents realizing information units between which native speakers might be expected to produce transition markers most often, and where the use of such markers is least likely to be disruptive of discourse coherence. In fact, the native speakers produced no such disfluencies in locations 5 and 6, between preposition phrases or between a preposition phrase and an adverb phrase, and they were significantly less likely to produce transition markers in the combined positions 1-6 than were the late immersion students ($p < .01$) and less likely at a marginal level of significance ($p < .10$) to do so than the early immersion and extended French students, too. Nonetheless, the native speakers' use of transition markers in combined positions 1-6 ($\bar{X} = 3.70$) was indeed very much higher than their use of such markers in all twelve other positions within phrases ($\bar{X} = 0.37$), where disfluencies can be assumed to be more disruptive of discourse coherence. Moreover, when transition markers in positions 1-6 were calculated for each group as a proportion of transition markers in all positions (1-18), it was evident that the native speakers were proportionately more likely to use such markers in positions 1-6, and correspondingly less likely in positions 7-18, than were any of the learner groups. For the native speakers the mean proportion of use in positions 1-6 was 92%, significantly higher than for each of the learner groups, whose proportions ranged from a mean of 81% for the early immersion students to 72% for the extended French students and 66% for the late immersion students. Among the learner groups there was a marginally significant difference between the 81% mean of the early immersion students and the 66% of the late immersion students, providing one indication that the late immersion students, being proportionately more likely to use transition markers in within-phrase locations, were less fluent overall than the early immersion students.

As is clear from Table 8, all three learner groups were in fact significantly more likely ($p < .01$) than the native speakers to produce transition markers in the combined within-phrase locations of 7-18. The early immersion students were, at the same time, less likely overall to produce within-phrase transition markers than the late immersion students, although they did not differ significantly from the extended French students. This provides additional evidence that the late immersion students were less native-like than the early immersion group in overall fluency. Table 8 shows that there were few significant between-group differences for individual within-phrase locations. Any such specific differences should, in any event, be treated with caution, since each group may have produced different total numbers of specific within-phrase contexts. It is, however, relevant to note that those locations where significant differences between the learner groups and the native speakers emerged (locations 8, 12 and 13), all involved the use of transition markers prior to major open word classes -- verbs and nouns, suggesting that accessibility of vocabulary was an important factor affecting the fluency of the learner groups.

5:3 Summary of Fluency Findings

On most of the types of disfluency that were analysed, the three learner groups produced significantly more disfluencies than the native speakers but did not differ significantly from one another. In a few instances, however, in the area of transition markers and cut offs, there was evidence that the immersion students were more fluent than the other learner groups.

With respect to the location of transition markers, the findings indicate that, of the learner groups, only the late immersion students were significantly more likely than the native speakers to use transition markers at clause and phrase boundaries. Within phrases, however, all three learner groups were significantly more likely overall than the native speakers to use transition markers. The early immersion students were, at the same time, significantly less likely than the late immersion students to use within-phrase transition markers, although they did not differ in this regard from the extended French students. The fluency findings are further discussed in section 7 below.

6. SOCIOLINGUISTIC TEST RESULTS

The results of the sociolinguistic oral test, administered to the three learner groups but not to the native speakers in this study, are displayed in Table 9, p. 231. As can be seen from the total scores for the low situational variants and the high situational variants, all three groups appropriately used more formal markers in the high situational variants. There were, however, very few differences found between the early immersion students and the other learner groups, and those differences that emerged were only near-significant at the $p < .10$ level. The early immersion students used slightly more conditional verb forms in the high situational variants than did the late immersion students ($\overline{X} = .11$ versus .00, $p < .10$), resulting also in a greater difference score on this variable for the early immersion students ($p < .10$). On the other hand, the early immersion students, as expected (see p. 204), demonstrated less sensitivity to the appropriate use of second person forms _tu_ and _vous_ than did the other learner groups. All three groups sometimes used _vous_ inappropriately in the low situational variant, and _tu_ inappropriately in the high situational variant. As indicated by their lower difference score, however, the early immersion students were less likely ($p < .10$) than either of the other groups to distinguish between use of person in the high and low situational variants.

The findings on this test suggest that, from a sociolinguistic perspective, the early immersion students do not appear to have gained an advantage over the other learner groups from their intensive exposure to French at a younger age.

7. DISCUSSION

It was hypothesized that, in general, the early immersion students in this study would be ahead of the other learner groups in the development of L2 proficiency, given the greater opportunity provided by their early start for obtaining needed L2 input and for making sustained oral communicative use of the L2. The findings on verb use, fluency, and sociolinguistic proficiency are now considered in relation to this hypothesis.

7:1 Verb Use

The findings in the area of verb use indicated some advantage, as predicted, but also some unexpected lack of advantage for the early immersion students.

A partial explanation for the mixed findings could be that the late immersion and extended French students have benefitted on certain variables both from early exposure to written input and from a more code-focussed introduction to the French verb system, made possible by their greater cognitive maturity on entry to their respective programs. From our knowledge of the type of audiolingual L2 texts that the late entry students have likely been exposed to, we can assume that there will have been more emphasis in the early stages on practising the kinds of formal features that these students did relatively well on, such as number agreement, past tense auxiliary use, and irregular past participles, and less emphasis on developing control of the more subtle semantic distinctions of the imparfait and conditional tenses, where the early immersion students were much further ahead. This suggested interpretation of some of the findings needs, of course, to be verified by classroom observation in the late entry programs.

Drawing on our current knowledge of the L2 development of early total immersion students (e.g. Harley and Swain 1984) and the learning activities and language use patterns evident in classrooms at the elementary level (Chapter 5), we can at the same time suggest some reasons why the much longer L2 experience of these students may not have led to a more consistent advantage over the late entry students.

At the early elementary level when these students begin their immersion program, an explicit focus on form is obviously much less feasible than for older students. To a large extent the learners are expected to — and in fact do — induce rules and features of the target language system from the context-embedded, largely content-focussed oral input provided by the teacher. However, because virtually all the students have the same mother tongue and because the natural content-focussed classroom talk in the L2 appears to have a number of limitations, the students' L2 development may proceed slowly once an initial stage has been reached where the students can satisfactorily master the content and can make themselves understood in the classroom context. Among the observed limitations of content-focussed classroom talk, for example, are that it does not necessarily provide learners with a great deal of exposure to some problematic L2 forms, nor does it automatically encourage the productive use of such forms by students. In the area of time distinctions, for instance, classroom observations in early immersion classrooms at grades 3 and 6 indicate that teachers spontaneously use mainly present tense and imperative verb forms when talking to the students, and that there is a limited amount of sustained talk by the students themselves (Chapter 5 and Appendix A to Chapter 5). Even when a seemingly 'natural' context for the oral use of past tenses arises in, for example, a history lesson, good content teaching strategies may involve making the issues more immediate to the students and thus give rise to very little oral use of past tenses by the teacher and none by the students (see example cited in Appendix D to Chapter 5).

A need for deliberately planned activities in the early immersion classroom context, designed to focus the L2 input and promote practice in the meaningful productive (oral and written) use of problematic L2 forms by students, has previously been expressed, and a study was carried out to test this hypothesis at the grade 6 level (Chapter 6). The results of that study led to the conclusion that in addition to focussed input and opportunities for student output, more corrective feedback from teachers would be beneficial in connection with such activities in order to provide students with increased incentive and needed information to advance their proficiency in specific linguistic domains. As is indicated from the classroom observations in grade 6 (Chapter 5), while focussing on content goals, teachers do not generally simultaneously provide regular correction of errors in the code. In short, without additional activities with a focus on the accurate, productive use of the code, the students may develop non-native patterns of use, strongly influenced by their shared mother tongue (Harley and Swain 1984), which become an established part of their mutually comprehensible interlanguage.

While the present study has not involved a detailed analysis of the kinds of errors made by the early immersion versus the late entry students, it is relevant to note that the early immersion students in grade 10 still appear to be making some of the same errors in the verb system that have been found in the performance of grade 1 early immersion students (Harley 1986): for example, in number and person agreement (les autres va aller -- 11105; son père vas être mouillé -- 11111)[4], in auxiliary use (j'ai allé -- 11109; elle a venu -- 12301), in the placement and form of object pronouns (elle va amener moi --12302; appeler elle sur le téléphone -- 11111) and in the use of verb vocabulary (j'étais à peu près 7 ans -- 11106; ça regarde comme ils descendent une colline -- 11108). Although such errors do not appear to be as frequent as in earlier grades (Harley and Swain 1984), their persistence at the grade 10 level suggests that some interlanguage fossilization has taken place. It may be hypothesized that more attention to these L2 features needs to be paid at a relatively early stage of the language acquisition process, in order to prevent their widespread establishment and continued use. How best to present and provide adequate practice in the use of these features in an age-appropriate way at the early elementary school level remains a major issue that requires more in-depth study and experimentation.

The lack of advantage on the part of the early immersion students in some areas of the verb system is not only due to errors. In the area of vocabulary use, for example, the early immersion students' well-developed strategic competence in the use of high coverage verbs appears to enable them to 'get by' without necessarily producing errors in their descriptions of the picture cartoons in the interview context. As noted elsewhere in Chapter 1, early immersion students in grades 5 and 6 still tend to rely largely on high coverage verbs that are congruent with their English L1 to get their meaning across. These findings, together with the observation that vocabulary teaching in the classroom focusses heavily on meaning and interpretation of words encountered in reading materials (Chapter 5), suggest that there is room for more emphasis in the early immersion context on ensuring that students make the move from general comprehension of words in context to the productive use of a richer, more varied vocabulary in their speech and writing.

The fact that the early immersion students in this study did substantially better than the late immersion and extended French students on the progressive use of the imparfait and on the hypothetical use of the conditional is interesting in light of the classroom observation finding that very little use was made of the relevant forms and functions in the natural talk of grade 3 and grade 6 early immersion teachers (Appendix A to Chapter 5). We may hypothesize that increasing amounts of written input and deliberate teaching of the use of these relatively marked L2 features have had a beneficial effect before the students encountered any strong need to express these functions with any great frequency and before any widespread patterns of non-native oral use had become established. That simple frequency in the naturally occurring oral L2 input in the immersion context is not necessarily a determining factor in its accuracy of use by the students is also suggested by another, contrasting finding in this study. The early immersion students were no more native-like on the second person plural variable than the late immersion students, despite classroom observation evidence from grade 6 that teachers make considerable use of plural vous-forms in their natural talk to the class (Chapter 5). An important element for the early immersion students in this finding could be a lack of sufficient practice in meaningful productive use of the plural forms. Surprisingly, the early immersion students appear to make less use of the appropriately colloquial on + verb stem in the first person plural context than the late immersion students, even though early immersion use in grade 1 tends to be more native-like on this variable (Harley 1986). The fact that some early immersion students in the present study were using the more formal, literary nous + -ons suggests that a later influence from written text may be reinforced by congruence with English where a distinction is made in the first person plural pronoun. It appears to be no coincidence that non-use of plural vous-forms is also congruent with English.

7:2 Fluency

There are several indications in the fluency findings that the early immersion students are overall more fluent than the other learner groups, as predicted on the basis of their greater experience in using the L2 for oral communication. From Table 7, p. 229, it is evident that there is a generally non-significant tendency for the early immersion students to produce fewer disfluencies than the other learner groups, except in the case of unfilled pauses. Most important, it seems, is that the early immersion students display advantages in the area of within-clause disfluencies, producing significantly fewer cut-offs than the extended French students, near-significantly fewer within-word cut-offs than the late immersion students, and significantly fewer transition markers in within-clause locations than the late immersion students. Perhaps even greater advantages in the area of fluency might be predicted, if more opportunities were provided in the early immersion context for sustained talk by students (Chapter 5). The tendency of all three learner groups to produce transition markers before verbs and nouns reinforces the suggestion made in the discussion of the verb use findings that more attention to the development of students' productive vocabulary would be beneficial. Lack of accessible vocabulary appears to be an impediment to the production of fluent speech.

7:3 Sociolinguistic Proficiency

In keeping with previous findings by Lapkin and Swain (1984a, b; Appendix C to Chapter 5), the early immersion students in the present study were not significantly superior to the late entry students on the sociolinguistic oral test[5] and tended to do less well than the later entry students on the formal/informal second person distinction, both with respect to use of tu and vous in the sociolinguistic test and with respect to 'polite' second person agreement in the verb (Question 11 of the interview). The proposed explanation of these findings is similar to that put forward in relation to verb use. The early immersion students, at least at the elementary level, appear to receive minimal sociolinguistically relevant input and to have little opportunity in their classroom context to rehearse different social roles (see also Swain and Lapkin 1986). Over time, it appears that a pattern of consistent use of tu becomes established that is congruent with English and generally appropriate in the classroom context. In contrast, late entry students have probably had explicit instruction at an early stage concerning the appropriate use of vous and tu.

In line with the verb use findings, the early immersion students were near-significantly superior to the late immersion students in the use of the conditional, in this case to express the notion of attenuation in formal situations. Once again, it may be suggested that the early immersion students have over time received more code-focussed input on the conditional that could have given them a slight advantage in its use in this 'optional' context too.

8. CONCLUSIONS

The findings of this study indicate that there are some advantages in oral L2 proficiency for an early start in an immersion program as opposed to a late entry immersion or extended French program. However, these advantages appear less widespread than might be anticipated given the much longer exposure to the L2 that the early immersion students have had. Of particular note is the comparable performance on many variables of the extended French students, whose exposure to the L2 has at no time been higher than about 40% of their school program. The findings give rise to a number of suggested implications for L2 curriculum in the early immersion context, including an early and continuing emphasis on providing carefully planned, age-appropriate focussed input in persistent problem areas of the L2 system and increased opportunities for meaningful, productive use of the relevant distinctions.

Footnotes

1 Twelve students were originally selected, but one student was absent at the time of oral testing.

2 The learner sample for the "Second Language Maintenance at the Secondary School Level" project consisted of one early immersion class and two extended French classes in the Toronto Board of Education, and two early immersion and two late immersion classes in the Carleton Board of Education. From each class, a stratified sample of eight students had been selected for oral testing. The remaining students represented the pool from which a sample for the present study was selected.

3 Slight differences on some of the native speaker scores compared with those reported in Harley (1986) reflect the fact that medians are calculated by a different method in SPSS-X from that earlier used in SPSS. In addition, the 'past' variable in the present study was based on a smaller set of interview questions, and an error found in the native speaker score for lexis was corrected.

4 In this study, five-digit numbers refer to specific students.

5 Lapkin and Swain did, however, note a tendency for early immersion students to be doing better than late entry students.

References

Genesee, F. Bilingual education of majority-language children: The immersion experiments in review. **Applied Psycholinguistics**, 1983, 4 (1), 1-46.

Harley, B. Second language proficiency and classroom treatment in early French immersion. Paper presented at the FIPLV/Eurocentres Symposium on Error in Foreign Language Learning, University of London, September 1985.

Harley, B. **Age in Second Language Acquisition.** Clevedon, Avon: Multilingual Matters, 1986.

Harley, B. and Swain, M. The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper and A.P.R. Howatt (eds.) **Interlanguage**. Edinburgh: Edinburgh University Press, 1984.

Hulstijn, J.H. Communication skills of young second language learners: Issues in coding speech data from face-to-face conversations. Toronto: Modern Language Centre, O.I.S.E., 1983.

Lapkin, S. and Swain, M. Second language maintenance at the secondary school level: Final report to the Carleton Board of Education. Toronto: Modern Language Centre, O.I.S.E., 1984a.

Lapkin, S. and Swain, M. Second language maintenance at the secondary school level: Final report to the Toronto Board of Education. Toronto: Modern Language Centre, O.I.S.E., 1984b.

Morrison, F. and Pawley, C. Evaluation of the Second Language Learning (French) Programs in the Schools of the Ottawa and Carleton Boards of Education, Vol. 1: French Proficiency of Immersion Students at the Grade 12 Level. Toronto: Ministry of Education, Ontario, 1986.

Olynyk, M., Sankoff, D. and d'Anglejan, A. Second language fluency and the subjective evaluation of officer cadets in a military college. **Studies in Second Language Acquisition**, 1983, 5 (2), 213-249.

Swain, M. Some inadequacies of typical content teaching for language teaching. Plenary paper presented at TESL Canada '87, Vancouver, B.C., March 1987.

Swain, M. and Lapkin, S. Immersion French in secondary schools: The 'goods' and the 'bads'. **Contact**, 1986, 5 (3), 2-9.

Table 1

Sample Characteristics

	Early Immersion	Late Immersion	Extended French	Native Speaker
Mean age (yrs:mths)	15:10	15:10	15:10	15:9
Mean I.Q.*	114.9	117.2	117.6	110.0
Region:				
Ottawa	9	12	0	0
Toronto	3	0	11	0
Quebec	0	0	0	12
Sex:				
Male	5	4	4	6
Female	7	8	7	6
Languages used at home:				
English	8	12	9	0
English, French sometimes	3	0	0	0
English, other	1	0	0	0
Other	0	0	1	0
English, other, French sometimes	0	0	1	0
French	0	0	0	12

*Learners' I.Q. was measured in grade 9 by the Otis Lennon Mental Abilities Test, Intermediate Form J. There were no significant differences between the learner groups. Native speakers' I.Q. was measured by the Épreuve d'Habileté Mentale Otis-Lennon, Niveau supérieur, Forme J, at the grade 10 level.

Table 2

Inter-rater Reliability of Percentage Verb Variables and Total Verb Scores for Learners (n = 12) and Native Speakers (n = 12)

	Learners			Native speakers		
	n	t	r	n	t	r
Percentage variables:						
Future	12	1.39	.99***	12	-1.55	.88***
Periphrastic future	12	0.45	.97***	12	-1.17	.73**
Past	12	1.37	.80**	0		
Imparfait	12	-0.71	.97***	12	-0.48	.98***
Conditional	12	-1.45	.80**	12	-1.36	.77**
First person plural	12	0.34	.91***	12	1.32	.92***
First person on	12	-1.09	.96***	12	1.32	.92***
Second person plural	11	-1.00	.99***	3	(1)	
Second person polite	9	1.00	.92***	11	(1)	
Third person plural	12	-1.79	.81**	12	(1)	
Pronominal verbs	12	-1.53	.85***	12	-1.00	(2)
Total scores:						
Clitic pronoun complements	12	0.56	.96***	12	0.00	.85***
Lexis	12	0.67	.89***	12	-0.59	.81**

*	$P < .05$
**	$P < .01$
***	$P < .001$

(1) All students scored 100% by both scorers.
(2) All students scored 100%, except for one by one scorer.

Table 3

Comparisons between Groups on Total Scores Using Mann-Whitney

	MEANS				MEDIANS				SIGNIFICANCE OF COMPARISONS					
	Early (N = 12)	Late (N = 12)	Extended (N = 12)	Native (N = 12)	Early (N = 12)	Late (N = 12)	Extended (N = 12)	Native (N = 12)	Early /Late	Early /Extended	Late /Extended	Early /Native	Late /Native	Extended /Native
Future	4.5	4.7	3.0	5.4	4.0	5.0	3.0	5.5						**
Periphrastic future	1.5	1.8	1.5	4.3	1.5	1.0	1.0	4.0				**	**	**
Past	4.0	5.7	4.7	12.0	3.5	5.0	3.0	11.0				**	*	*
Imparfait	2.6	1.4	0.4	6.8	2.7	1.0	0.0	7.0	*	***	+	***	***	***
Conditional	1.8	0.5	0.3	6.6	2.0	0.0	0.0	5.5	*	**		***	***	***
1st person plural	1.8	3.8	3.4	8.1	1.5	4.0	2.6	8.0	+	*		***	***	***
1st person on	1.0	2.8	1.2	8.1	0.6	3.0	1.0	8.0				***	***	***
2nd person plural	0.5	1.2	0.5	1.0	0.0	1.0	0.0	0.0						
2nd person polite	0.8	1.0	1.0	1.4	1.0	1.0	1.0	1.5				*		
3rd person plural	2.9	4.3	3.8	5.3	3.0	3.0	2.0	5.0				*		+
Clitic pronoun complements	4.3	3.1	1.5	5.3	4.5	2.0	1.0	5.0		**		+	*	***
Pronominal verbs	2.0	1.4	1.5	4.4	1.5	1.0	1.0	4.0				*	**	**
Lexis	26.2	30.8	27.3	37.8	26.0	31.5	25.0	36.5	+			***	*	***

* $p < .05$, 2 tailed
** $p < .01$, 2 tailed
*** $p < .001$, 2 tailed
+ $p < .05$, 1 tailed

Table 4

Comparisons between Groups on Percentage Variables Using Mann-Whitney

	MEANS				MEDIANS				SIGNIFICANCE OF COMPARISONS					
	Early (N = 12)	Late (N = 12)	Extended (N = 12)	Native (N = 12)	Early (N = 12)	Late (N = 12)	Extended (N = 12)	Native (N = 12)	Early /Late	Early /Extended	Late /Extended	Early /Native	Late /Native	Extended /Native
Future	81.6	71.1	51.1	91.4	91.7	80.4	50.0	100.0		*			+	***
Periphrastic future	30.9	28.8	28.8	72.3	31.0	20.8	25.0	82.9				**	**	**
Past	67.9	66.6	53.6	95.1	75.0	78.1	50.0	100.0				**	**	**
Imparfait	54.3	13.7	4.1	63.9	63.3	13.4	0.0	70.7	**	***	*		***	***
Conditional	48.8	12.5	5.1	94.6	55.0	0.0	0.0	100.0	*	**		**	***	***
1st person plural	45.8	82.3	83.8	93.4	50.0	100.0	100.0	100.0	*	*		**		
1st person on	25.7	56.1	23.9	93.4	12.5	71.8	25.0	100.0	+		+	***	*	***
2nd person plural	(10) 43.3	(10) 65.3	(9) 40.7	(3) 100.0	(10) 16.7	(10) 100.0	(9) 0.0	(3) 100.0				+		+
2nd person polite	(11) 45.5	(11) 77.3	(9) 77.8	(11) 100.0	(11) 50.0	(11) 100.00	(9) 100.0	(11) 100.0	+	+		***	+	
3rd person plural	69.7	55.0	53.8	100.0	73.3	50.0	50.0	100.0				**	***	***
Pronominal verbs	70.8	46.8	54.5	95.8	83.3	37.5	50.0	100.0				*	**	**

* p < .05, 2 tailed
** p < .01, 2 tailed
*** p < .001, 2 tailed
+ p < .05, 1 tailed
Ns indicated in parentheses for 2nd person variables

808

807

Table 5

Minimum and Maximum Total Scores on Each Verb Variable

	Early Immersion (N = 12)		Late Immersion (N = 12)Min		Extended (N = 12)		Native Speaker (N = 12)	
	Min	Max	Min	Max	Min	Max	Min	Max
Future	2	9	1	9	0	7	2	9
Periphrastic Future	0	4	0	6	0	4	1	7
Past	0	10	0	13	0	13	2	21
Imparfait	0	6	0	6	0	2	4	9
Conditional	0	4	0	4	0	2	3	13
1st person plural	0	5	0	10	1	10	3	12
1st person on	0	5	0	7	0	3	3	12
2nd person plural	0	2	0	4	0	2	0	8
2nd person polite	0	2	0	2	0	3	0	2
3rd person plural	0	6	0	12	1	15	1	11
Clitic pronoun complements	0	10	0	8	0	3	3	8
Pronominal verbs	0	6	0	4	0	4	1	9
Lexis	19	33	19	39	20	40	31	46

Table 6

Minimum and Maximum Percentage Scores on Each Verb Variable

	Early Immersion (N = 12)		Late Immersion (N = 12)Min		Extended (N = 12)		Native Speaker (N = 12)	
	Min	Max	Min	Max	Min	Max	Min	Max
Future	20	100	17	100	0	100	60	100
Periphrastic Future	0	100	0	100	0	57	20	100
Past	0	100	0	100	0	100	77	100
Imparfait	0	100	0	38	0	20	27	100
Conditional	0	100	0	100	0	33	60	100
1st person plural	0	100	0	100	0	50	38	100
1st person on	0	100	0	100	50	100	38	100
2nd person plural	0	(10) 100	0	(10) 100	0	(9) 100	100	(3) 100
2nd person polite	0	(11) 100	0	(11) 100	0	(9) 100	100	(11) 100
3rd person plural	0	100	0	100	20	100	100	100
Pronominal verbs	0	100	0	100	0	100	50	100

*Numbers in brackets indicate numbers of students producing contexts for use of the given form. For these two variables, students are treated as "missing" if no context was produced, while on other variables a percentage score of zero was assigned if no contexts were produced.

Table 7

Type of Disfluency

	Mean per 100 words per student				Significance (2-tailed t-test)				
	Early	Late	Extended	Native	Early /Late	Early /Extended	Early /Native	Late /Native	Extended /Native
Progressive									
Repeats	3.45	4.23	4.67	1.18			**	**	**
'uh' etc.	7.61	10.68	7.82	4.07			*	**	*
Regressive									
Unfilled pauses	2.45	1.51	1.59	0.04			**	**	**
Repair conversions:									
(a) within word	0.59	0.66	0.50	0.08			**	**	*
(b) within string	0.87	1.02	1.05	0.22			*	*	**
Cut offs:									
(a) within word	0.41	1.49	1.40	0.17	+	**	**	*	**
(b) sound change	0.57	0.89	0.87	0.10			+	*	**
(c) structural change	1.77	1.32	1.97	0.73				+	**
Progressive	11.06	14.91	12.49	5.25			**	**	**
Repair conversions (a) + (b)	1.46	1.67	1.55	0.30			**	**	**
Cut offs (a) + (b) + (c)	2.75	3.70	4.24	1.00		*	**	**	**
Regressive	6.67	6.87	7.38	1.35			**	**	**
All disfluencies	17.73	21.78	19.87	6.60			**	**	**
Progressive/All	61%	67%	61%	82%			**	*	**

+ p < .10
* p < .05
** p < .01

811

812

Table 8

Location of Disfluencies

Location[a]	Mean per 100 words per student				Significance (2-tailed t-test)				
	Early	Late	Extended	Native	Early /Late	Early /Extended	Early /Native	Late /Native	Extended /Native
Between clauses or phrases									
1	3.39	4.50	2.95	2.66				*	
2	1.37	1.65	1.58	0.20			**	**	**
3	0.73	0.51	0.86	0.62					
4	0.31	0.20	0.32	0.22					
5	0.02	0.03	0.00	0.00					
6	0.00	0.09	0.02	0.00					
Within phrases									
7	0.02	0.20	0.14	0.01	+			+	
8	0.29	0.75	0.49	0.04	+		+	**	**
9	0.17	0.20	0.12	0.05					
10	0.00	0.08	0.02	0.00					
11	0.09	0.04	0.00	0.01					
12	1.14	1.35	0.88	0.07			*	**	**
13	0.04	0.46	0.16	0.02	*			*	+
14	0.00	0.22	0.10	0.10					
15	0.00	0.10	0.00	0.00					
16	0.05	0.19	0.09	0.02				+	
17	0.00	0.12	0.09	0.05					
18	0.00	0.00	0.00	0.00					
1-6	5.82	6.98	5.73	3.70			+	**	+
7-18	1.79	3.71	2.09	0.37	*		*	**	**
1-18	7.61	10.68	7.82	4.07			*	**	*
1-6/1-18	81%	66%	72%	92%	+		*	**	**

[a] See Appendix C, pp. 245-247, and Figure C2, pp.249-250, for the syntactic description of each of these numbered locations.

+ p < .10
* p < .05
** p < .01

813

814

Table 9

Results on Sociolinguistic Oral Test

Type of formal marker	$\overline{X}$ use of formal markers			Significance levels (2-tailed t-test)	
	Early Immersion	Late Immersion	Extended	Early /Late	Early /Extended
Low situational variant:					
Introduction	.06	.04	.05		
Interrogative	.64	.64	.68		
Conditional	.00	.00	.00		
Vocabulary	.07	.13	.08		
Attenuation	.14	.08	.12		
Politeness	.10	.15	.09		
Person	.48	.53	.37		
Total	1.48	1.57	1.38		
High situational variant:					
Introduction	.71	.69	.54		
Interrogative	.86	.94	.84		
Conditional	.11	.00	.05	+	
Vocabulary	.39	.26	.29		
Attenuation	.17	.17	.15		
Politeness	.25	.25	.22		
Person	.93	1.49	1.34		
Total	3.41	3.81	3.43		
Difference scores:					
Introduction	.65	.65	.49		
Interrogative	.22	.31	.16		
Conditional	.11	.00	.05	+	
Vocabulary	.32	.14	.22		
Attenuation	.03	.08	.03		
Politeness	.15	.10	.13		
Person	.45	.96	.97	+	+
Total	1.94	2.24	2.05		

+ $p < .10$

Appendix A

INTERVIEW SCHEDULE

1. Tu as quel âge?

2. C'est quand, ta fête/ton anniversaire?

3. Quand c'est ta fête, qu'est-ce que tu aimes manger le plus?

4. Qu'est-ce que tu sais préparer toi même (à la cuisine)?
4a. (Si rien) As-tu jamais aidé ta mère à faire un gâteau?

5. Est-ce que tu peux m'expliquer comment préparer ... (ce que l'élève a répondu à 4)?

6. Où est-ce que tu habites? Quel est ton adresse?

7. Est-ce que c'est loin d'ici?

8. Si tu étais en retard pour l'école, qu'est-ce que le directeur/professeur ferait? Qu'est-ce qu'il ferait si tu arrivais souvent en retard?

A. 9. Je vais te montrer une petite histoire en images. Il y a une bonne idée pour aider les gens qui risquent d'être en retard le matin. Veux-tu me raconter l'histoire de ce pauvre monsieur? Ça commence ici et continue par ici, tu vois.

10. Oui, c'est ça. Et qu'est-ce qu'il va faire maintenant, tu penses?

11. Imagine maintenant que tu rencontres ce monsieur dans la rue et que tu veux savoir quelle heure il est. Alors, qu'est-ce que tu lui dis? Tu veux être poli, n'est-ce pas, parce que tu ne le connais pas. Alors, moi, je fais le monsieur. Qu'est-ce que tu dis au monsieur?

12. D'habitude, qui est-ce qui te réveille le matin?
12a. (Si c'est pas la mère) C'est pas ta mère alors?

13. Qu'est-ce qu'elle faisait, ta mère, quant tu es parti pour l'école ce matin?

14. Est-ce que tu as un animal chez toi?
14a. (Si oui) Raconte-moi quelque chose que ton chien/chat/etc. a fait?

15. As-tu des frères ou des soeurs?

16. Est-ce qu'il(s)/elle(s) fait/font partie du programme d'immersion?
16a. (Si oui) En quelle année?
16b. (Si l'élève n'a pas de frères ou de soeurs) As-tu des amis qui habitent près de chez toi?

17. Qu'est-ce que vous aimez faire ensemble?
17a. (Si ce n'est pas dans la réponse à 17) A quels jeux jouez-vous?

18. Est-ce que tu peux m'expliquer comment jouer à ...? (ce que l'élève a mentionné en répondant à 17)

19. Qu'est-ce que tu feras ce soir après l'école?
19a. Et puis, qu'est-ce que tu feras? ... Et puis? (jusqu'à ce que l'élève se mette au lit)

20. Pourquoi est-ce qu'on doit se coucher de bonne heure pendant la semaine?

21. Tu regardes la télévision quelquefois? Quel est ton programme favori?

22. Raconte-moi ce qui s'est passé la dernière fois que tu as vu ...? (le programme mentionné)

23. Avez-vous déjà fait un voyage en famille, tous ensemble?
23a. (Si non) Avez-vous déjà fait une visite quelque part en classe?

24. Qu'est-ce que tu as aimé le mieux de ce que tu as vu?

25. Comment étai(en)t ...? (quelque chose que l'élève a mentionné)

26. Et ta famille, avez-vous des projets pour les vacances de Noel/d'été/de Pâques qui viennent?

27. Qu'est-ce que tu aimes le mieux, l'été ou l'hiver? Pourquoi?

28. L'hiver/l'été passé, qu'est-ce que tu faisais d'habitude les fins de semaine (le weekend)?

B. 29. J'ai ici une petite scène d'été. Voici un homme avec un appareil photo. Qu'est-ce qui se passe?
29a. Qu'est-ce que tu penses qu'il dit aux enfants?
29b. Pourquoi les enfants sont-ils contents à la fin?

C. 30. Voici une autre scène d'été. Qu'est-ce qui arrive ici?
30a. (Si plonger n'a pas été utilisé) Qu'est-ce qu'il fait ici (6e dessin)?
30b. Pourquoi est-ce qu'il y a tant de pieds dans le sable?
30c. Qu'est-ce que tous ces gens auraient dû faire avant de sauter?

D. 31. Bon, on va changer de saison, maintenant. On va regarder trois petites scènes d'hiver. Veux-tu me raconter l'histoire de cette image? (Montrez D).
31a. Pourquoi le Monsieur n'est pas content?

E. 32. Voici un accident de skidoo (moto-neige). Qu'est-ce qui se passe ici?

F. 33. Voici une autre sorte d'accident. Qu'est-ce qui est arrivé?
33a. Où est-ce qu'ils vont l'amener, tu penses?

34. As-tu déjà été à l'hôpital?
34a. (Si non) As-tu jamais eu une grande peur?
34b. (Si non) As-tu jamais fait un mauvais rêve?

35. Raconte-moi ce qui s'est passé.

36. Qu'est-ce que tu ferais, si tu avais beaucoup d'argent? Si tu gagnais la loterie, par exemple?

37. Et tes parents, qu'est-ce qu'ils feraient, eux?

38. Qu'est-ce que tu aimes le mieux faire à l'école?

39. Qu'est-ce que la classe faisait quand tu es sorti tout à l'heure?

40. Maintenant, imagine que tu es la maîtresse/le professeur et que la classe fait des bêtises. Qu'est-ce que tu dis aux élèves? Voici la classe devant toi. Alors, qu'est-ce que tu nous dis?

Appendix B

Scoring Procedures for Verb Use

Tim Tomlinson

The following scoring procedures were used in calculating learners' scores and in checking the already available native speaker scores for inter-rater reliability purposes.

1. MISSING DATA

A student's response was discarded as missing data when the interview question was either omitted entirely or unfairly rephrased by the interviewer. The latter case included changes which modelled or prompted a sought-after verb, or which jeopardized the chances of a particular form being produced. For example, on one occasion the question *Et qu'est-ce qu'il va faire maintenant, tu penses?* (question 10) was followed by *Penses tu qu'y va être en retard à six heures le matin, là?*, which unfairly models an additional verb in the periphrastic future (i.e. *va être*). Elsewhere, a follow-up question to *Qu'est-ce que tu feras ce soir après l'école?* (question 19) was reworded *Qu'est-ce que tu vas faire?*, thereby changing the context for the future.

2. GENERAL CRITERIA

The following eleven criteria for tallying verb forms were applied to all of the variables under analysis.

2:1 Parentheses

Material in parentheses in the transcripts (i.e. not clearly interpretable by the transcriber) was not included in the analysis.

2:2 Phonetic transcriptions

Verb forms appearing in phonetic transcriptions were omitted from analysis unless they could be interpreted from context as mispronunciations of recognizable forms, as with *il meurt* transcribed *il /m R/*. The utterance *un homme a /da/* was omitted in scoring one student's response to question 31 since it was unclear what verb was produced in a context logically requiring the verb DONNER.

2:3 False starts

Hesitation phenomena involving 'exact' repetitions of an immediately preceding verb form were eliminated from the tally. Thus, *c'est eum um c'est comme ça* counted as one verb only, whereas FAIRE, TOMBER, and CAUSER were each counted in *Ils ont faire tomber/Non ils ont causé un accident...* (21219 - question 32).[4]

2:4 Modelled items

If the interviewer modelled an item (often on request) which was not in the interview schedule, and which was repeated by the interviewee within the 'same' question, then such an item was not tallied in the analysis. For example, in question 9, one student was given SE LEVER and then used it in three different ways (*pour se lever*, *il n'est pas levé*, *il se lève*), none of which was counted for that particular question.

2:5 Confirmation checks by student

If, while groping for an elusive verb form, the interviewee actually used the correct form but sought confirmation from the interviewer as to its correctness, the initial and any subsequent occurrences of the item were tallied. For instance, *sourient* was counted in the following exchange for question 29:

Student - 32432	Interviewer
... il veut qu'ils uh...ils oh ... *sourient*? Ça c'est le mot?	
	Sourient oui.

2:6 Repetition of questions

Verb forms appearing in an 'exact' repetition of the interviewer's question were not included in the analysis. Thus, in responding with *Oh, oh qu'est-ce qu'ils auraient dû faire? Um ne pas être si haut* to question 30c, one student merely repeated a modelled form, *auraient dû faire*, without showing that he could use it in an original utterance.

2:7 Anglicisms, neologisms and paralinguistic means

English substitutions for French verbs, verb forms created by the student which exist in neither the English nor the French lexicons (e.g., *il graffe l'auto*: 21114 -question 31), or paralinguistic methods of conveying an action were counted as verb contexts and included for calculating base percentage scores where relevant.

2:8 Auxiliary - ALLER

ALLER used in the present tense with third person reference, together with an infinitive, was scored as an auxiliary in the periphrastic future unless it was clear from the context that ALLER was used as lexical verb. Thus, in *il va aller prendre une douche*, *va* would count as an auxiliary with respect to the lexical verb *aller*.

2:9 Semi-auxiliary FAIRE

Semi-auxiliary (causative) use of FAIRE was counted as a separate verb form from an immediately following verb. See 2:3 above for an example.

2:10 Auxiliaries ÊTRE and AVOIR

Forms of ÊTRE and AVOIR when followed by a verb form were counted as auxiliaries and not as separate verbs. This means that certain constructions which could be interpreted as ÊTRE + complement (e.g., *c'est fini*) were treated as Auxiliary + Verb. This criterion applies as well to verbs expressed in the passive voice, as in this response to question 29: *il va être ... uh frappé de l'eau* (21220), where *va* and *être* were considered auxiliaries. When a form of AVOIR or ÊTRE was left dangling as a result of a false start, it was not tallied owing to the impossibility of determining whether it should be interpreted as an auxiliary or a lexical verb.

3. CRITERIA SPECIFIC TO INDIVIDUAL VARIABLES

The following scoring procedures pertain only to the respective variables under which they are listed. For each variable, students were assigned a 'total score' based on a straightforward count of appropriately realized target items. These items did not necessarily have to be grammatically correct in all respects to be counted as appropriate in context. Formal and semantic deviations which were included in the total score are indicated for each variable. Where a 'percentage base' was also calculated, it consisted normally of all obligatory contexts the student provided for the variable in question. This obviously included items tallied in the total score. Where applicable, a 'percentage score' was calculated for each student by dividing the total score by the percentage base.

3:1 Past time

Several questions in the interview were designed to elicit past tense forms in the context of narratives. For the purposes of this analysis, only data appearing in response to questions 23, 34, 34a, 34b and 35 were considered since answers to these questions most consistently elicited past time reference. In cases where students were invited to tell more than one narrative for questions 34, 34a and 34b (usually because little or no response was given initially), only the 'first proper narrative' was scored.

The total score included the following forms produced in past time contexts:

(1) appropriately realized items in the passé composé, the imparfait and other past tenses,

(2) unambiguous errors of number in the auxiliary and of gender in the past participle of the passé composé (e.g., les enfants a souri),

(3) use of auxiliary AVOIR in composed past tenses of the verb TOMBER, a common feature of native speaker discourse (e.g., elle a retombé: 43544 - question 34).

The percentage base consisted of all verb contexts with past time reference, including:

(1) occurrences of the historic present (which native speakers occasionally use to add immediacy to past events in their narratives),

(2) errors in the past tense interpretable as the wrong auxiliary, as in on a retourné (21219 - question 23) where est was required,

(3) incorrect past participle forms (e.g., mon père a prend: 12301 - question 34),

(4) ambiguous forms consisting of verb stem + /e/ in contexts calling for the passé composé: (e.g., il regard/e/: 11105 - question 34; je juste jou/e/: 32330 - question 34).

3:2 Future time

Scoring for this variable was based on data elicited in response to questions 10 and 19.

The total score included:

(1) instances of periphrastic future, simple future, conditional, and elliptical infinitive forms denoting future time (e.g., Prendre une douche..je pense:11109-question 10).

(2) unambiguous number and person errors in future forms, including variants of je vais typical of informal Canadian French (e.g., M'a appeler mes amis: 43538 - question 19; Je vas faire mes devoirs: 43539 - question 19).

The percentage base consisted of all future time contexts, including the use of present tense to refer to future events. Unless habitual meaning was clearly indicated by adverbial means (e.g., Euh, généralement, j'écoute un petit peu la télévision: QS03 - question 19), the verb form was assumed to have occurred in a future time context.

Finally, in a series of infinitives directly following a single instance of auxiliary ALLER, only the first infinitive was scored as a periphrastic future, the others being recorded as infinitives. Thus, in Là, il va aller faire sa toilette, pis déjeuner (question 10), déjeuner was counted as an elliptical infinitive, not as a periphrastic future form.

3:3 Periphrastic future (ALLER + infinitive)

This variable was scored using the same questions and procedures as for Future time, the only difference being that the total score consisted of the periphrastic future alone.

3:4 Imparfait

This variable was scored on the basis of answers to three questions, two of which were designed to elicit reference to incomplete actions in the past (13 and 39), the other reference to habitual past actions (28).

The total score was obtained by summing the number of imparfait forms produced in response to the three questions. The percentage base consisted of all finite verb forms produced in the responses, except for verbs used in expressions such as je ne sais pas, je ne me rappelle plus, c'est tout, je pense, ça fait que, etc. which clearly did not refer to the past.

3:5 Conditional

Scoring of this variable was based on answers to questions 36 and 37, each of which involved remotely possible hypothetical events.

The total score was calculated by summing all appropriately realized conditional forms in hypothetical contexts supplied by the student. The percentage base included these forms as well as incorrect forms in obligatory conditional contexts. The most common errors included the use of an infinitive, a present tense form or a periphrastic future form to express hypothetical modality.

3:6 First person plural

Reference to the actions, states, etc. of the speaker plus others was elicited through questions 17, 17a, 23, 23a and 26. The total score included forms that agreed in number and person with whichever subject pronoun (on or nous) the student had selected, regardless of:

(1) errors in tense,
(2) wrong auxiliary,
(3) errors in accompanying infinitives,
(4) wrong past participle,
(5) minor errors in lexical verb stem (e.g., nous viennons: 32326 - question 26).

The percentage base included, in addition to the above items, number agreement errors and elliptical infinitives in contexts for first person plural.

3:7 On as first person plural

This variable was treated in the same way as the preceding one, except that the total score consisted simply of forms agreeing in (unmarked) number and person with the subject pronoun on.

3:8 Second person plural

A total score for this variable was arrived at by summing the second person plural forms used in response to question 40. Minor errors in the verb (e.g., s't'arrêtez!: 11104) as well as regional variants (e.g., Assoyez-vous: 21217) were also included.

The percentage base consisted of all contexts for second person, including

(1) errors in the verb stem and/or inflection (e.g., taites-vous instead of taisez-vous: 21220),
(2) second person singular forms (e.g., ferme ta bouche: 11109).

Note that some non-native speakers occasionally mixed direct and indirect discourse within the same clause. Owing to difficulties of interpretation, such instances were not included in the total or percentage scores. Answers in which only indirect discourse was used were considered as missing data.

3:9 Polite second person

Data for this variable was elicited in question 11 which required making a polite request to an adult stranger. The total score for politeness was based on the number of second person 'polite' (plural) verb forms produced by the student. The percentage base consisted of all second person contexts supplied by the student and thus included the sociolinguistic error of addressing an adult stranger using 'familiar' singular instead of 'polite' (plural) second person forms (e.g., Excuse-moi, ... tu sais le temps?: 12301).

3:10 Third person plural

Several questions (29, 29b, 30, 30b, 32 and 33) required a description of picture cartoons depicting the actions of persons other than the speaker and the addressee at the interview.

The total score consisted of:

(1) appropriately realized, phonologically distinct third person plural forms. These included cases where an obligatory liaison was actually made between a pronoun-subject (i.e. *ils* or *elles*) and a verb (phonologically distinct or not) beginning with a vowel or mute 'h' (e.g., *ils avaient*; *ils ont*). Where the transcriber judged that such a liaison was not made (e.g., *il(s) avai(en)t* and *ils/arrivent*: 43538 - questions 30 and 32), the item was discarded from analysis, as were other non phonologically distinct forms such as *ils lui portent* (11109 - question 33),

(2) phonologically distinct third person plural forms containing errors in past participles (e.g., *ils ont prend*: 21220 - question 33),

(3) phonologically distinct third person plural forms containing errors involving the use of *ont* and *sont*, particularly when used as auxiliaries in the context of *passé composé* (e.g., *ils ont allé*: 21218 - question 30).

The percentage base consisted of all obligatory contexts for phonologically distinct third person verb forms, including:

(1) number errors in main or auxiliary verbs (e.g., *ils veut* and *les enfants a souri*: 21104 - question 29)

(2) errors in choice of mood (e.g., *ils sachent* instead of *ils savent*: 11111 - question 29).

3:11 Clitic pronoun complements (excluding pronominal verbs)

The analysis of non-subject, clitic pronouns was based on questions 5, 9, 10, 12, 31, 31a, 33 and 33a. Only a total score was calculated, which consisted of non-reflexive, preverbal pronoun complements, including:

(1) gender errors in the pronouns,

(2) errors in verb number, person or tense (e.g., *je la mis*: 21217 - question 5; *je le serve*: 32330 - question 5),

(3) syntactic errors in negative structures, such as *il la ne fait pas* (11112 - question 31).

Omitted from the total score were:

(1) errors in pronoun case (e.g., *ils lui emportent*: 11107 -question 33),

(2) errors involving placement of the clitic pronoun after the verb, as in on metter le dans le four: 32327 - question 5).

3:12 Pronominal verbs

Questions 9, 10 and 12 were used in the scoring of pronominal verbs. The total score consisted of all contextually appropriate pronominal verbs accurately expressed in terms of reflexive pronoun and, where relevant in the context of the passé composé, auxiliary ÊTRE. Also included were:

(1) minor errors in the reflexive pronoun, such as se instead of s' in the infinitive s'habiller (21219 - question 10),

(2) errors in tense, past participle and infinitives of pronominal verbs just as long as a recognized verb stem form was produced.

The percentage base consisted of all contexts provided for pronominal verbs and included:

(1) overgeneralization of reflexive pronouns, as in le machine se casse le nez de la personne (32325 - question 9),

(2) omission of reflexive pronouns (e.g., il doit réveiller: 21219 - question 10; il va brosser ses dents instead of il va se brosser les dents: 21219 - question 10).

(3) incorrect pronouns in a pronominal verb context (e.g., pour lui préparer where se was required: 21220 - question 10).

3:13 Lexis

For this variable, the questions involving picture descriptions (9, 10, 29, 30, 31, 32, 33 and associated questions) were selected since they provided circumscribed topics and thus contexts which would be most likely to elicit similar lexical items from the native speakers. A single score was calculated consisting of the number of different lexical verbs (i.e., verb types) produced across the above contexts. Hence, only the first occurrence of a lexical verb was tallied regardless of the number or variety of individual forms (i.e., verb tokens) produced thereafter. In cases where a verb appeared in both its pronominalized and non-pronominalized forms, it was counted as a single item unless the lexical meanings of the two forms are entirely distinct. Thus, réveiller and se réveiller would be considered two tokens of the same verb type, whereas entendre and s'entendre would be regarded as separate lexical verb types. The lexical score also included:

(1) syntactic and morphological errors in verb constructions, especially the omission or overgeneralization of reflexive pronouns in pronominal verbs (e.g., se tomber: 21224 - question 30),

(2) past participles used with verbal meaning (e.g., Rendu à six heures, il dort encore: 43538 - question 9) and present participles,

(3) marginal lexical errors which do not impede comprehension; for example, gratter instead of égratigner to express the notion of scratching a car with a shovel (question 31), or voir instead of vérifier in voir le profondeur de l'eau 12303 - question 30); decisions as to what constitute marginal semantic errors were, of course, a matter of individual interpretation.

The total score for lexis excluded:

(1) more serious lexical errors such as son alarme va instead of sonne (21221 - question 9), non-existent verb forms such as déprendre (32325 - question 31) and anglicisms such as squeezer (43548 - question 9),

(2) nouns or adjectives derived from verbs (e.g., un sourire in question 29),

(3) auxiliaries ALLER, AVOIR and ÊTRE (see 2:8 and 2:10).

Appendix C

Procedures for Fluency Scoring

Patrick Conteh

1. INTERVIEW QUESTIONS USED FOR FLUENCY SCORING

Q.9 Je vais te montrer une petite histoire en images. Il y a une bonne idée pour aider les gens qui risquent d'être en retard le matin. Veux-tu me raconter l'histoire de ce pauvre monsieur? Ça commence ici et continue par ici, tu vois?

Q.18 Est-ce que tu peux m'expliquer comment jouer à ..... (name of some sport or game mentioned earlier)

Q.22 Raconte-moi ce qui s'est passé la dernière fois que tu as regardé ce programme.

2. WORD COUNT

(a) Every lexical item within an utterance was counted separately:

e.g. qu'est-ce qu'il y a (= 7 words)
1 2 3 4 5 6 7

(b) Repair initiators such as 'uh', 'em', etc., unfinished words in cut-offs, or sound change segments are not included in the word count.

(c) Lexical items such as 'okay', 'bon' etc. (which indicate 'acknowledge' and 'accept' of turn taking, or surprise) are included in the word count.

(d) Repeats of words are included in the word count as many times as there is recurrence.

(e) All borrowed words as well as titles of films, books etc. and proper names are included in the word count.

e.g. Fun with Dick and Jane uh
1 2 3 4 5

était ... était un p ... peu em
6 7 8 9

un peu [ ] scary (= 12 words)
10 11 12

3. CRITERIA FOR SCORING HESITATION OR DISFLUENCY MARKERS

A. Progressive Repairs

(1) Repeats. Any recurrence in a sequence (or with some intervening 'uh', 'em', etc.) of some lexical element or structure (other than the disfluency marker) is counted as a 'Repeat' of that element or structure

e.g. Il ..... il (= 1 Repeat of 'il');

il ..... il il	(= 2 Repeats of 'il');
il um il	(= 1 Repeat of 'il');
il um il um il	(= 2 Repeats of 'il')..

A recurrence of any of the other disfluency markers is scored as the 2nd, 3rd, etc. occurrence of the disfluency marker.

e.g.	il uh uh	(= 2 occurrences of 'uh');
	il um il um	(= 1 Repeat of 'il', 2 occurrences of 'um')

(2) uh's etc. in isolation. Repair initiators such as 'uh' 'um' 'ben' etc.. which occur between words or phrases.

e.g.	il um veut	(= 1 uh in isolation)
	il uh .. uh il	(= 2 uh's, 1 repeat of 'il')

B. Regressive Repairs

(3) Unfilled pauses. Any pause of two or more seconds within an utterance is counted. Latencies until the onset of the subjects' turn are not counted.

(4) Repair conversion within a word. A word-internal phenomenon in which some segment(s) within a word is/are changed or repaired and in which there is the repetition of some previous item(s) in the string.

e.g. ce serait plus intesserant intéressant;
il le il l'écrit.

(5) Repair conversion within an expression. Change or repair within an expression involving the addition or replacement of some lexical item(s) and the possible repetition of some previous item(s) or structure(s) within the expression.

e.g. il le il la voit /il le la voit;
avec si avec tant de personnes;
je veux être une peut-être une vétérinaire.

(6) Cut-offs within a word. A word-internal phenomenon in which there is a break, or some hesitation phenomenon within a word such that the first part of the word is incomplete but is later completed. This may involve a glottal stop.

e.g. les oi ..... les oiseaux;
les ois ..... oiseaux;
les en ,,,,, les enfants.

Or it may involve some other word-internal break followed by a different lexical item.

e.g. les enf ..... garçons.

(7) Cut-offs: sound change. The intrusion of some unwanted, isolated and meaningless sound segment that is abandoned in favour of the required meaningful lexical item.

e.g. il voulait [ ] partir.

(8) Cut-offs: structural change. A break within an expression or a sentential structure involving the abandoning and replacement of an unfinished expression or sentential structure in favour of another expression or sentential structure.

e.g. il y avait ..... les trois garçons voulaient .....

4. CRITERIA FOR SCORING CONTEXTS FOR HESITATION MARKERS

(1) $[_{S}\ \#(X)\_\_\ ]$: clause-initial position, with or without a preceding X, where X represents some category such as Adverb, Adverbial phrase (AdvP), preposition, prepositional phrase (PP), conjunction etc.

e.g. "puis em il se lève"
"quelques minutes après um il s'est levé"
"il s'est lavé et um il est allé ....."

(2) $[_{S}\ NP\_\_(VP)]$: where within a clause the marker occurs between an NP and a VP, if any.

e.g. "ça um se réveille l'homme"
"c'est un soldat qui um ..... "

(3) $[_{VP}\ V(AdvP\_\_(NP)]$: where a marker occurs after a verb or after a verb and an AdvP and followed (or not) by an NP.

e.g. "l'appareil lui pince em ....."
"l'appareil lui pince em le nez "
"ça prend quand même euh une force"

(4) $[_{VP}\ V(NP)\_\_ \left\{\begin{matrix} PP \\ AdvP \end{matrix}\right\}]$: where a marker occurs between a Verb and either a PP or an AdvP that may or may not be preceded by an NP.

e.g. "il pousse le ballon um dans le filet"
"on jouait euh oh d'habitude"

(5) $[_{VP}\ PP\_\_AdvP]$: between a PP and an AdvP

e.g. "avec uh presque toujours ..... "
e.g "il va dans la rue em très vite"

(6) $[_{VP}\ (PP)\_\_PP]$: between two PPs or before PP

e.g. "sans l'appareil um dans son nez "

(7) $[_{VP}\ Aux\_\_(Part)]$: between the auxiliary verb and a following participle, if any

e.g. "un appareil lui a um pincé le nez"

(8) [V_(Inf)]: between a matrix verb and an infinitival complement (if any):
VP

e.g. "il voulait em se lever"
"il doit um ..... "

(9) [Prep_(V)]: between a preposition and a Verb (if any)
VP

e.g. "pour euh ..... mis le 'puck'"

(10) [Adv_V]: between an Adverb and a Verb
VP

e.g. "il vite em met son .....

(11) [Cl_(V)]: between a clitic and a Verb (if any)
VP

e.g. "il le em pousse"

(12) [X_(Adj)(N)]: between X and a nominal element (if any) preceded by a
VP qualifying adjective (if any); where X represents some determiner
(article, possessive etc. .....)

e.g. "deux um personnes ..... "
"le um pauvre monsieur ..... "

(13) [Adj_(N)]: between a qualifying Adjective and a Noun (if any). Numeral adjectives
VP are considered as determiners.

e.g. "les petites um filles ....."
"toutes la même euh ..... "

(14) [(NP)_NP]: between a preceding NP (if any) and a following NP.

e.g. "le petit orphelin euh Albert ..... "

(15) [XN_PP]: between a Noun (preceded by some determiner, possessive etc. X) and a
NP PP modifier within NP.

e.g. "le monsieur um dans la photo ..... "

(16) [P_(NP)]: where within a PP some hesitation marker occurs after a preposition or
PP between a preposition and a following NP.

e.g. "pour em .... pour le ....."
"avec uh ..... quat/quatre personnes ..... "

(17) [Deg_Adj]: where within an Adjectival phrase a marker occurs between the degree
AdjP marker and the adjective.

e.g. "était pas mal euh scrappé ..... "
"ni trop euh faible ..... "

(18) [Deg__Adv]: where within an Adverbial phrase a marker occurs between the degree
AdvP marker and the adverb:

e.g "très um lentement il ..... "

Figure C1

Scoring Sheet: Type of Disfluency

Student #	Progressive		Regressive						Word Count
			3 unfilled pauses	Repair conversions		Cut Offs			
	1 repeats	2 uh in isolation		4 within word	5 within ex-pression or string	6 within word	7 sound change	8 struc-tural change	

Figure C2

Fluency Scoring Sheet: Location of Transition Markers

Student#	1 [#(X)__] S	2 [NP__(VP)] S	3 [V(AdvP) __(NP)] VP	4 [V(NP)__ PP] VP AdvP	5 [(PP)__AdvP] VP	6 [(PP)__PP] VP	7 [Aux__(Part)] VP	8 [V__(Inf)] VP

Figure C2 (continued)

Fluency Scoring Sheet: Location of Transition Markers

Student#	9 [Prep__(V)] VP	10 [Adv__V] VP	11 [Cl__(V)] VP	12 [X__(Adj)(N)] NP	13 [Adj__(N)] NP	14 [(NP)__NP] NP	15 [XN__PP] NP	16 [P__(NP)] PP	17 [Deg__Adj] AdjP	18 [Deg__AdV] AdvP

Appendix D

Oral Sociolinguistic Scoring Scheme

In scoring this test, students' reponses to each picture stimulus were analysed with respect to the presence or absence of formal markers. Seven categories of formal markers were identified, and for most of these one point was given if a student used a formal marker. Use of more than one formal marker in the same category did not give rise to extra points. The seven categories, with examples, are listed below.

(1) Introduction

The following types of utterance openers were considered to be formal markers of politeness. One point was assigned for the use of one or more of these introductions:

pardonnez-moi
excusez-moi
je m'excuse
pardon
s'il vous plaît
monsieur/madame

(2) Interrogative construction

Use of an interrogative construction of the following nature was counted as a formal marker. One point was given for the use of one or more such constructions:

Inversion (voulez-vous ..., peux-tu ...)
Est-ce que ...
Je veux savoir si ...

Note: Use of an intonation question (e.g. tu veux t'asseoir près de moi) was not considered a formal marker.

(3) Conditional verb form

One point was given for the use of the conditional verb form:

auriez-vous ...
je voudrais ...

(4) Formal vocabulary and extra explanatory information

One point was given if a response contained any formal vocabulary and/or extra explanatory information associated with the response. Formal vocabulary included, for example:

prêter, absent, offrir, siège, utiliser, car, discuter

Examples of additional explanatory information were:

pour mieux voir
je veux lire ...

(5) Attenuations

1 point was given for markers of attenuation such as un peu, moins:

je veux lire un peu
faire moins de bruit

(6) Additional politeness markers (outside introduction)

1 point was assigned if one or more of the following types of politeness markers were used during a response, outside the introduction:

s'il vous plaît/s'il te plaît
monsieur/madame
merci
pardon
excusez-moi/excuse-moi

(7) Use of person

This category was scored somewhat differently from the others, in order to distinguish between those students who consistently used formal vous (or votre, vos) and those who used it mixed with informal tu (or ton, ta, tes). Two points were given for consistent use of vous-forms in a response, while one point was given for mixed use of vous- and tu-forms, and no points were given for consistent use of tu-forms. The implicit use of vous in an expression such as excusez-moi was counted in the scoring, but s'il vous plaît/s'il te plaît was not included in scoring for this category.

E.g. (a) Excusez-moi, est-ce que je peux emprunter votre règle?
(consistent vous = 2 points)

(b) Excusez-moi, est-ce je peux emprunter ta règle?
(mixed = 1 point)

(c) Excuse-moi, est-ce que je peux emprunter ta règle?
(consistent tu = 0 points)

Table D1

Initial Inter-rater Agreements on Scoring of Each Formal Marker for Each Item on 20 Cases

Number of Agreements

Item	Introduction	Use of Person	Verb Form	Verb Tense	Lexicon	Attenuation	Politeness Markers
Low status variants:							
Request 1	20	17	20	20	17	20	19
Request 2	20	20	20	20	10	15	20
Offer 1	20	16	20	20	11	15	19
Offer 2	20	19	20	20	19	18	20
Complain 1	19	n.a.*	20	20	15	14	19
Complain 2	20	17	20	20	18	20	20
High status variants:							
Request 1	20	16	20	20	16	19	19
Request 2	20	20	20	20	10	13	20
Offer 1	19	20	18	20	9	11	18
Offer 2	20	19	19	19	19	20	20
Complain 1	20	n.a.*	19	19	12	14	19
Complain 2	19	17	19	19	16	17	19

*Use of person (tu/vous) was not scored on "Complain 1" because of the presence of two addressees in the slide.